Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAA
GGAAAGTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACTTC
AGTGACTGAGGCCAGGCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTT
TCATAGACACTGATGACACGTTTACGCAAAATAGAAATTTGAGGAGAAACGCCTGG
GCCTTCGGAAAGGAGTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAA
CTAACTAATGTATACTATTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGC
AGCAGGAAAGCTTTGGTTAATTTGGAAATGGATGATAGCATTAAAATAACAGAAGC
GCCTCCAGGTCTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC
CACTAAGCCTTTTGATCCCTTTGGAAGCAAAGAACTTTCCTTCCCTGGGGTGAAGAC
TCTCCTCAGAAGATTTCCTGTCTCTGCCTATGTTACAAGAGGAATCAAAACCAAGAC
AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAGTCAGCGGCTTGTTGATCTCCC
TCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGGCCTGTCCTCGCCGCTGTG
CCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCGGTACCTGACCTCCATCCCAG
ATGGCATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAACAGCCTTACTAGAT
TGACAGAAAACGACTTTGATGGCCTGAGCAAACTGGAGTTACTCATGCTGCACAGT
AATGGCATTCACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAGTCCTTGCAGGTC
TTAAAAATGAGCTATAACAAAGTCCAAATCATTCGGAAGGATACTTTCTACGGACTC
GGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG
GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACA
AAGCTCCATCCAGACACATTTGTCTCATTAAGCTATCTCCAGATATTTAAAACCTCTT
TCATTAAGTACCTGTTCTTGTCTGATAACTTCCTGACCTCCCTCCCAAAAGAAATGGT
CTCCTACATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGA
CTGCCATTTAAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAAT
GCAAGAAAGACAGAAGCTCTTCCAGTCCTCAGCAATGTCCCCTTTGCATGAACCCCA
GGATCTCTAAAGGCAGACCCTTTGCTATGGTACCATCTGGAGCTTTCCTATGTACAA
AGCCAACCATTGATCCATCACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAAT
GGATCTGCCTCCACCTCACCTCAAGATTTCATAGAACCCTTTGGCTCCTTGTCTTTGA
ACATGACANANNTNTCTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA
TCAAGGACATCACCAACTGCATTCACTGAAGAAAATGACTACATCATGCTAAATGC
GTCATTTTCCACAAATCTTGTGTGCAGTGTAGATTATAATCACATCCAGCCAGTGTG
GCAACTTCTGGCTTTATACAGTGACTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT
TACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGGTGGCTCTTAGGCCTGAAGA
CATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTTTTTGGTTCCAACAAGA
AAAAATTGTCTTGCAGCTGAACAGAACTGCCACCACACTTAGCACATTACAGATCCA
GTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAGCGGAGAGAC
TCAAATGGACCATGATCCTGATGATGAACAATCCCAAACTGGAACGCACTGTCCTGG
TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTCACCTCACTTGG
AATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAAATGGGAAGTTGGAACTGCAGATGGCTGACAGCTTTGA
TGCAGGTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATA
CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC
ACACAGTGGTTACGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAG
ATGCTTCTATTAGCTGGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAG
ACAGGCAAATTCTTAACAATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGAT
CAAGGTCATTACCAATGTGTGGCTGCCAACCCATCAGGGGCCGACTTTTCCAGTTTT
AAAGTTTCAGTTCAAAAGAAAGGCCAAAGGATGGTTGAGCATGACAGGGAGGCAG
GTGGATCTGGACTTGGAGAACCCAACTCCAGTGTTTCCCTTAAGCAGCCAGCATCTT
TGAAACTCTCTGCATCAGCTTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGTG
TACATAGGAAGAACAAACATAGAGACTTAATACATCGGCGGCGTGGGGATTCCACG
CTCCGGCGATTCAGGGAGCATAGGAGGCAGCTCCCTCTCTCTGCTCGGAGAATTGAC
CCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATTCTGTGCCAAAAAA
GCAAGAAAATACCACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCGTGGAACTCA
CTGACGAGGAAAAGGATGCCTCTGGCATGATTCCTCCAGATGAAGAATTCATGGTTC
TGAAAACTAAGGCTTCTGGTGTCCCAGGAAGGTCACCAACTGCTGACTCTGGACCAG
TAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
CACAAACACTACAATCTGAGCACCTTCCTGATTTCAAATTATTTAGTGTAACAAACG
GTACAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACA
ACCAACCAAAACCCAATCATTATCTTTCCATCAGTAGCTGAAATTCGAGATTCTGCT
CAGGCAGGAAGAGCATCTTCCCAAAGTGCACACCCTGTAACAGGGGGAAACATGGC
TACCTATGGCCATACCAACACATATAGTAGCTTTACCAGCAAAGCCAGTACAGTCTT
GCAGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA
GCAGACCTAGCAGTAGTGACATCTCTTCTCACACTACTGCAGACCCTAGCTTCTCCA
GTCACCCTTCAGGTTCACACACCACTGCCTCGTCTTTATTTCACATTCCTAGAAACAA
CAATACAGGTAACTTCCCCTTGTCCAGGCACTTGGGAAGAGAGAGGACAATTTGGA
GCAGAGGGAGAGTTAAAAACCCACATAGAACCCCAGTTCTCCGACGGCATAGACAC
AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTCC
AGCCACAGAGTACCCTGGGATGTGCCACACATGTCCTTCCGCAGAGGGGCTCACAG
TGGCTACTGCAGCACTGTCAGTTCCAAGTTCATCCCACAGTGCCCTCCCCAAAACTA
ATAATGTTGGGGTCATAGCAGAAGAGTCTACCACTGTGGTCAAGAAACCACTGTTAC
TATTTAAGGACAAACAAAATGTAGATATTGAGATAATAACAACCACTACAAAATAT
TCCGGAGGGGAAAGTAACCACGTGATTCCTACGGAAGCAAGCATGACTTCTGCTCC
AACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTGGTCACCTGAGCATGCC
TGGGACCATCCAAACTGGGAAAGATTCAGTGGAAACAACACCACTTCCCAGCCCCC
TCAGCACACCCTCAATACCAACAAGCACAAAATTCTCAAAGAGGAAAACTCCCTTG
CACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGAATCCATATCA
ATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAAATAGCTCCTCTTTTACC
CACAGGTCAGAGTTCCCCCTCAGATTCTACAACTCTCTTGACAAGTCCGCCACCAGC
TCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAGTGCTTC
CTAGCACCATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCCCA
CAGTGACAAGTCCTACTGCTACTGCATCTGTCATTATGTCTGAAACCCAACGAACAA
GATCCAAAGAAGCAAAAGACCAAATAAAGGGGCCTCGGAAGAACAGAAACAACGC
AAACACCACCCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGC
TGATACCCCCTTGGCTTTCAGTCATTCCCCACGACAAGATGATGGTGGAAATGTAAG
TGCAGTTGCTTATCACTCAACAACCTCTCTTCTGGCCATAACTGAACTGTTTGAGAA
GTACACCCAGACTTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAAT
CACAGGAGAGTACCACAGTGAAAAGAGCCTCAGACACACCACCACCACTCCTCAGC
AGTGGGGCGCCCCCAGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTT
ACAGACAGCAAAGTCACATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCAT
ATATGAATCTTCAAGGCACAATACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCC
CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTTCCATCCAC
TTCTGTGCCAGCACTAAGGGTAGATAAACCACAGAATTCTAAATGGAAGCCCTCTCC
CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAAGGGCA
AAAGGCCAGCAGTAAGCATGTCCCCCCACCTCAGCCTTCCAGAGGCCAGCACTCAT
GCCTCACACTGGAATACACAGAAGCATGCAGAAÁAGAGTGTTTTTGATAAGAAACC
TGGTCAAAACCCAACTTCCAAACATCTGCCTTACGTCTCTCTACCTAAGACTCTATTG
AAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTACAGTTCCAGCTAATTC
AGACGTTTTTCTTCCTTGTGAGGCTGTTGGAGACCCACTGCCCATCATCCACTGGACC
AGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCCGGTTCCACGT
GCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTGGACAGTA
CCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTTGTCTGTG
GTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACTTT
GGAAGTACTGTGGAACTAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTC
CTGGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGG
TCTGGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTG
GTTTTTACAAGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCACTGTTGGTTAAGA
TACAAGTCATCACAGCTCCCCCTGTCATTATAGAGCAAAAGAGGCAAGCCATCGTTG
GGGTTTTAGGTGGAAGTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTA
GTGTTCACTGGGTCCTTTATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATT
CCAGATTTTTCTTGTATCCAAATGGAACTCTGTATATAAGAAGCATCGCTCCTTCAGT
GAGGGGCACTTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAG
TGATTCTTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAAACTGCCTCTCAG
AAATGGACTGAGGTGAATTTGGGTGAGAAATTACTACTGAACTGCTCAGCTACTGG
GGATCCAAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCATCGACCAGT
GGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGGTGGTTGGG
TCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAAGAAACAAAATGGG
AGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCAAAATTGAACA
GAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTGACTGCA

AGGCCTCTGGCTCCCCTGTGCCTGAGGTATCCTGGAGTTTGCCTGATGGGACAGTGC
TCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCTT
TTCCACAATGGAACCTTGTATTTCAACAACGTTGGGATGGCAGAGGAAGGAGATTAT
ATCTGCTCTGCCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGT
TCTAACAGCCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGG
CTGGAGAAACAGCTGTCCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTA
TTTTGGTTGCTGCCTTCCAACAATGTCATTTCATTCTCCAATGACAGGTTCACATTTC
ATGCCAATAGAACTTTGTCCATCCATAAAGTGAAACCACTTGACTCTGGGGACTATG
TGTGCGTAGCTCAGAATCCTAGTGGGGATGACACTAAGACATACAAACTGGACATT
GTCTCTAAACCTCCATTAATCAATGGCCTGTATGCAAACAAGACTGTTATTAAAGCC
ACAGCCATTCGGCACTCCAAAAAATACTTTGACTGCAGAGCAGATGGGATCCCATCT
TCCCAGGTCACGTGGATTATGCCAGGCAATATTTTCCTCCCAGCTCCATACTTTGGA
AGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAGATGAGGAACATCCGGCTTTC
TGACTCTGCGGACTTCACCTGTGTGGTTCGGAGCGAGGGAGGAGAGAGTGTGTTGGT
AGTGCAGTTAGAAGTCCTAGAAATGCTGAGAAGACCAACATTCAGAAACCCATTCA
ACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACTGCTCTGTGGAT
GGGAACCCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGTTTGCTAAC
AGACCACACAATTCCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAA
GCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCTA
CATCGAGAAACTCATCCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGA
ACCAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTGTCTG
ATGGGATCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGAC
AGGCCTCAAGTGGATGGAAAATACATACTGCATGAAAATGGCACGCTGGTCATCAA
AGCAACAACAGCTCACGACCAAGGAAATTATATCTGTAGGGCTCAAAACAGTGTTG
GCCAGGCAGTTATTAGCGTGTCAGTGATGGTTGTGGCCTACCCTCCCCGAATCATAA
ACTACCTACCCAGGAACATGCTCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGT
GTGGCCTTGGGAATCCCCAAGCCAAAAGTCACCTGGGAGACGCCAAGACACTCCCT
GCTCTCAAAAGCAACAGCAAGAAAACCCCATAGAAGTGAGATGCTTCACCCACAAG
GTACGCTGGTCATTCAGAATCTCCAAACCTCGGATTCCGGAGTCTATAAGTGCAGAG
CTCAGAACCTACTTGGGACTGATTACGCAACAACTTACATCCAGGTACTCTGACAGG
AAGGGGGAGACTAAAATTCAACAGAAGTCCACATCCACAGGGTTTATTTTTTGGAA
GAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGAATACATTTACAGT
ATTAAATTTACAATGGACATGCGATGAGACTTGTAAATGAAAGCATTGTGAACTGA
AACCGAGTCTCTGTGGATCTCAAAGCAAACTCTTAACTTAAGGCACTTTGATTTTGC
CAACAAATAATAACAAACATTAAGAGAAAAAAAATGATCCACTACGAAATAACAAAC
GGCTAATGCACCTGAATTCTCAGTAAAAAGACCTTTCTCTCGCTAACAGTTGCCAGC
TGCCTCGTGTCTGTTTCCTACCAATGTCACAAACATCGCACACAGGGTGAATGGAGT
CAACGGGAAAGATTAAGTTTGCGGTCTGTGTAAATCTCAATGTACAAATATTCTGTC
NCTGGTTTATAAACATTTTGATAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAA (SEQ ID NO: 1)

Figure 2

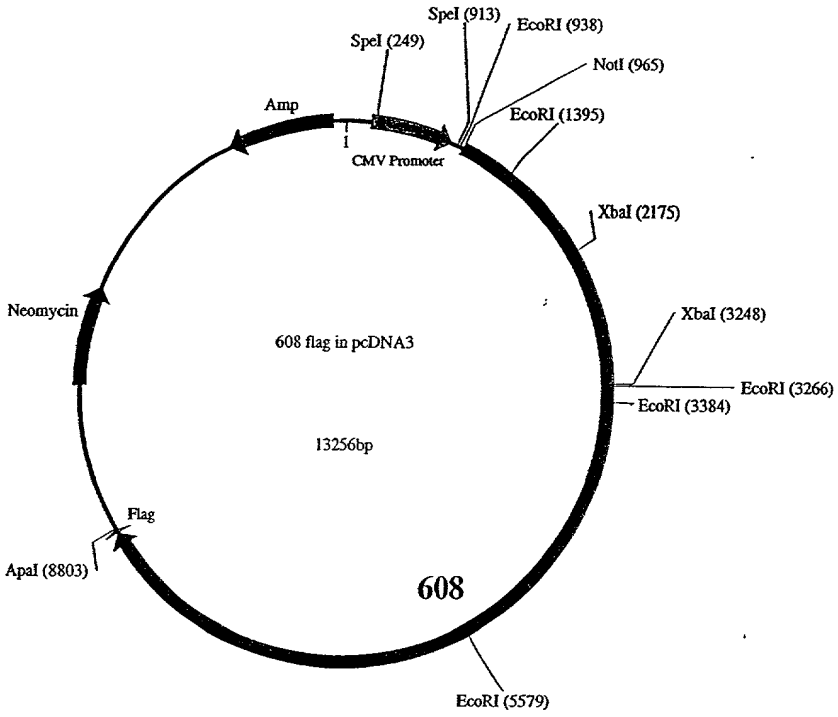

SpeI (913)
EcoRI (938)
SpeI (249)
NotI (965)
Amp
EcoRI (1395)
1
CMV Promoter
XbaI (2175)
Neomycin
XbaI (3248)
608 flag in pcDNA3
EcoRI (3266)
EcoRI (3384)
13256bp
Flag
ApaI (8803)
608
EcoRI (5579)

Figure 3

MQVRGREVSGLLISLTAVCLVVTPGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
RINLGYNSLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQI
FKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESLYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNGS
ASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFSTNL
VCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADVR
ADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNNPK
LERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQMA
DSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKDQGHYQCVAANPSGADFSSFKV
SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKLSASALTGSEAGKQVSGVHRK
NKHRDLIHRRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV
KPVPLAVPLVELTDEEKDASGMIPPDEEFMVLKTKASGVPGRSPTADSGPVNHGFMTSI
ASGTEVSTVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIIFPSVAE
IRDSAQAGRASSQSAHPVTGGNMATYGHTNTYSSFTSKASTVLQPINPTESYGPQIPITGV
SRPSSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGRV
KNPHRTPVLRRHRHRTVRPAIKGPANKNVSQVPATEYPGMCHTCPSAEGLTVATAALS
VPSSSHSALPKTNNVGVIAEESTTVVKKPLLLFKDKQNVDIEIITTTTKYSGGESNHVIPTE
ASMTSAPTSVSLGKSPVDNSGHLSMPGTIQTGKDSVETTPLPSPLSTPSIPTSTKFSKRKTP
LHQIFVNNQKKEGMLKNPYQFGLQKNPAAKLPKIAPLLPTGQSSPSDSTTLLTSPPPALST
TMAATQNKGTEVVSGARSLSAGKKQPFTNSSPVLPSTISKRSNTLNFLSTETPTVTSPTAT
ASVIMSETQRTRSKEAKDQIKGPRKNRNNANTTPRQVSGYSAYSALTTADTPLAFSHSP
RQDDGGNVSAVAYHSTTSLLAITELFEKYTQTLGNTTALETTLLSKSQESTTVKRASDTP
PPLLSSGAPPVPTPSPPPFTKGVVTDSKVTSAFQMTSNRVVTIYESSRHNTDLQQPSAEAS
PNPEIITGTTDSPSNLFPSTSVPALRVDKPQNSKWKPSPWPEHKYQLKSYSETIEKGKRPA
VSMSPHLSLPEASTHASHWNTQKHAEKSVFDKKPGQNPTSKHLPYVSLPKTLLKKPRIIG
GKAASFTVPANSDVFLPCEAVGDPLPIIHWTRVSSGXEISQGTQKSRFHVLPNGTLSIQRV
SIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILDRHVKEITVHFGSTVELKCRVEGMP
RPTVSWILANQTVVSETAKGSRKVWVTPDGTLIIYNLSLYDRGFYKCVASNPSGQDSLL
VKIQVITAPPVIIEQKRQAIVGVLGGSLKLPCTAKGTPQPSVHWVLYDGTELKPLQLTHS
RFFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVVILTVEEGETIPRIETASQKWTEVN
LGEKLLLNCSATGDPKPRIIWRLPSKAVIDQWHRMGSRIHVYPNGSLVVGSVTEKDAGD
YLCVARNKMGDDLVLMHVRLRLTPAKIEQKQYFKKQVLHGKDFQVDCKASGSPVPEV
SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFNNVGMAEEGDYICSAQNTLGK
DEMKVHLTVLTAIPRIRQSYKTTMRLRAGETAVLDCEVTGEPKPNVFWLLPSNNVISFS
NDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDIVSKPPLINGLYANKT
VIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMRNIRLS

DSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIAQAGKPVALNCSVDGNP
PPEITWILPDGTQFANRPHNSPYLMAGNGSLILYKATRNKSGKYRCAARNKVGYIEKLIL
LEIGQKPVILTYEPGMVKSVSGEPLSLHCVSDGIPKPNVKWTTPGGHVIDRPQVDGKYIL
HENGTLVIKATTAHDQGNYICRAQNSVGQAVISVSVMVVAYPPRIINYLPRNMLRRTGE
AMQLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGTLVIQNLQTSDSGVY
KCRAQNLLGTDYATTYIQVL (SEQ ID NO: 2)

Figure 4

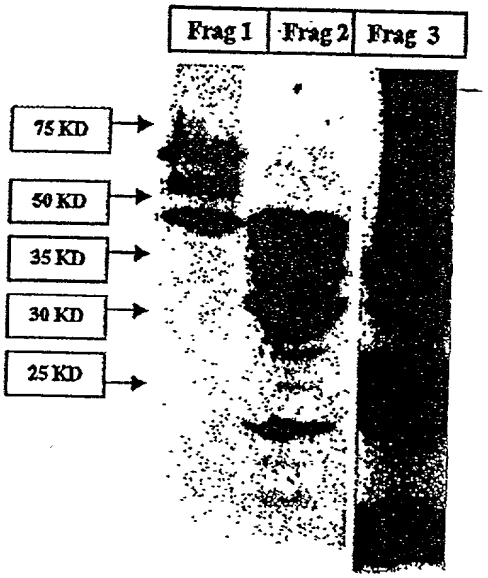


Figure 5

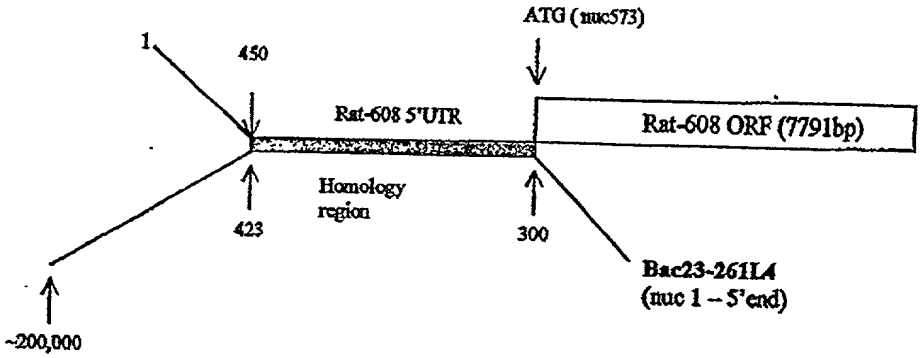


Figure 6

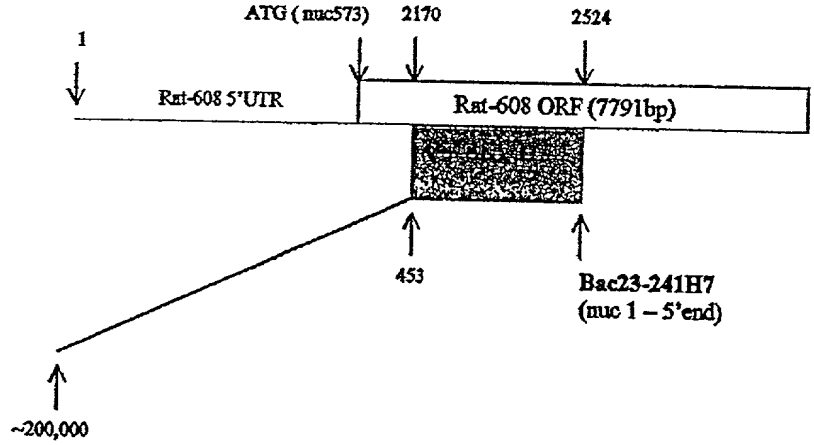
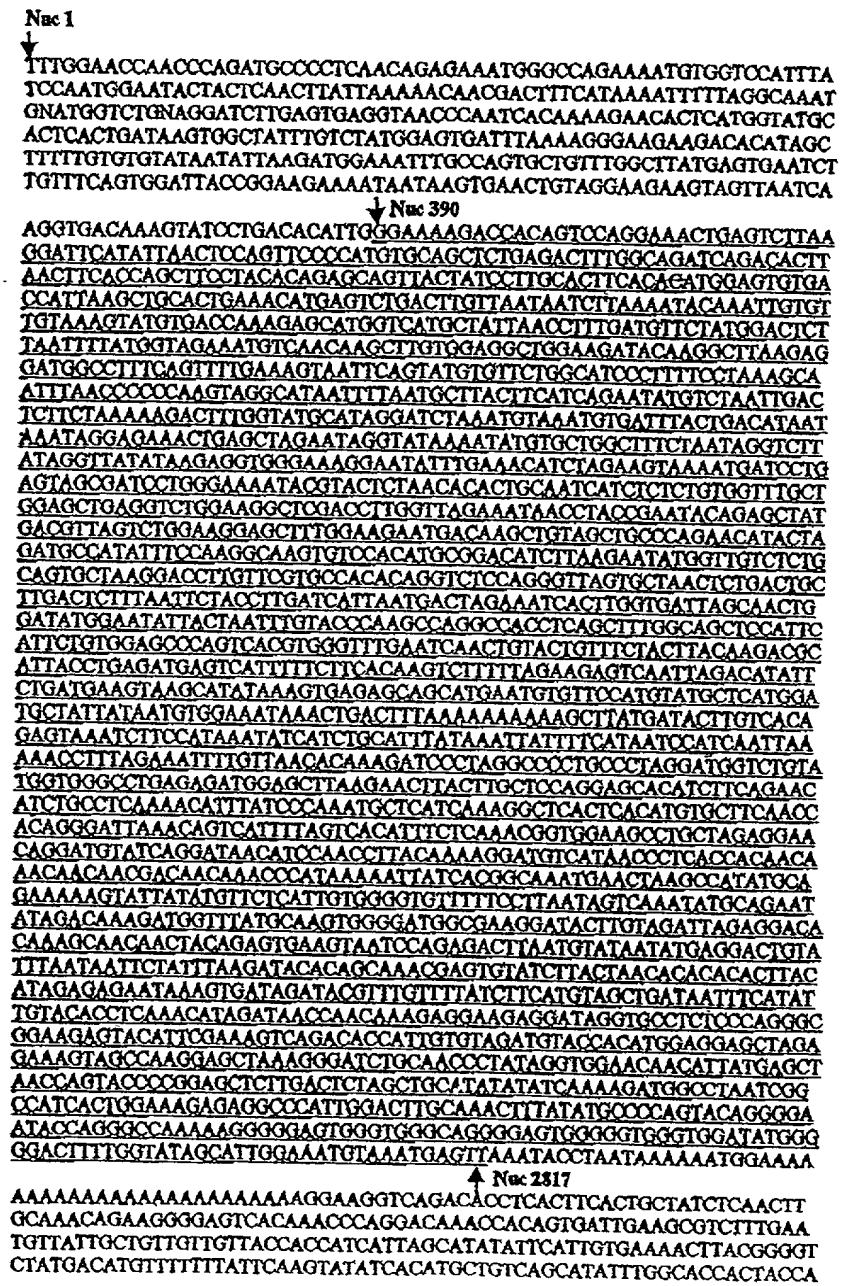
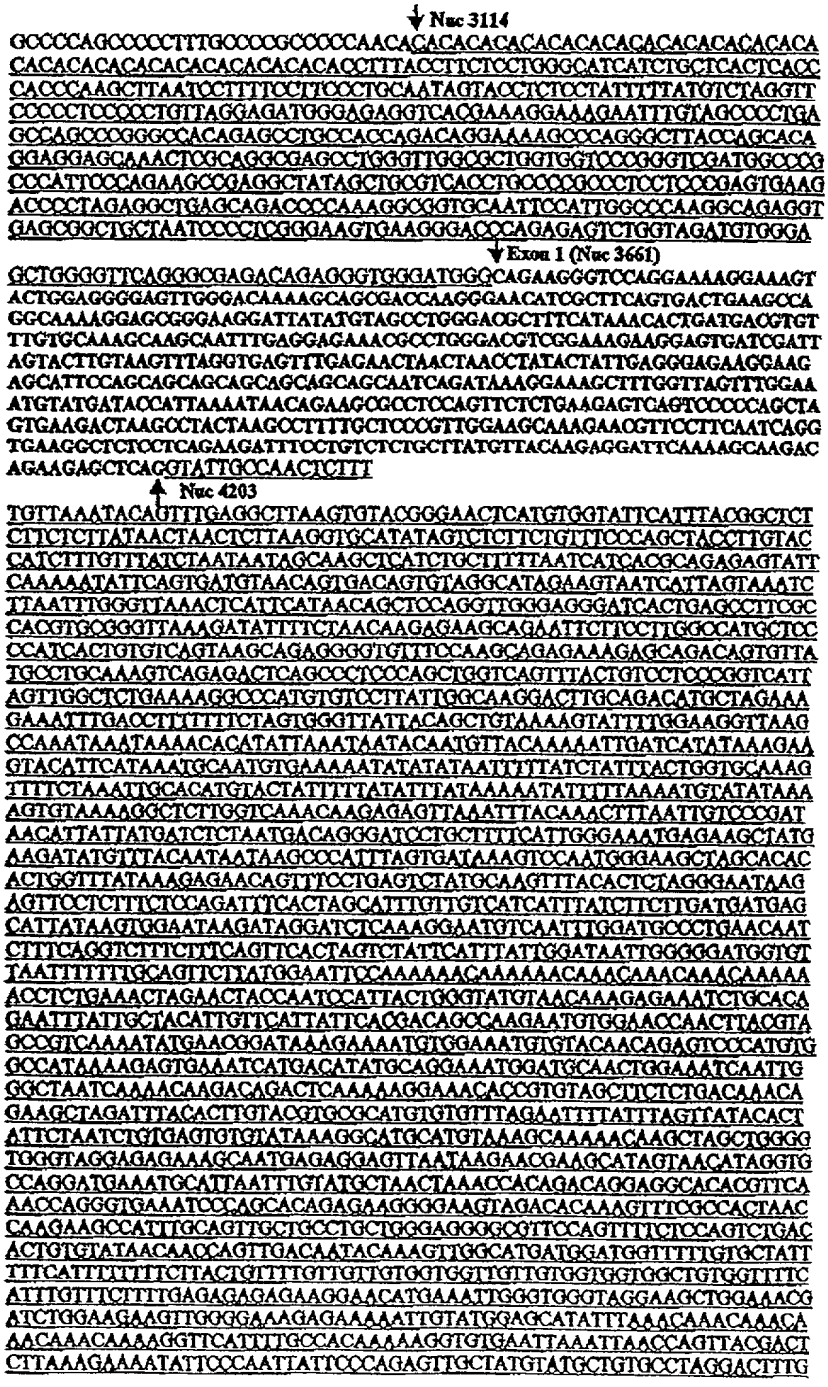
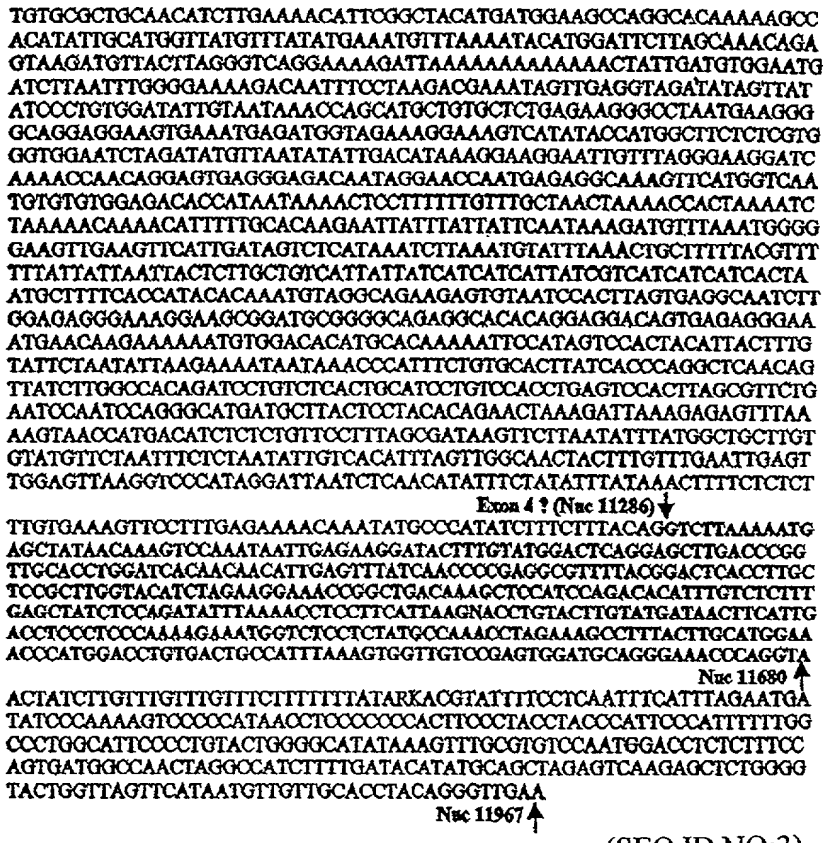


Figure 7




Nuc 3114
Exon 1 (Nuc 3661)
Nuc 4203


Exon 2 (Nuc 6559)
Nuc 6755
Exon 3 (Nuc 8089)
Nuc 8218



(SEQ ID NO:3)

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

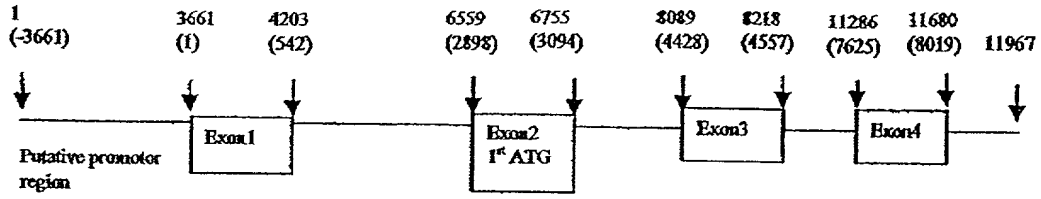
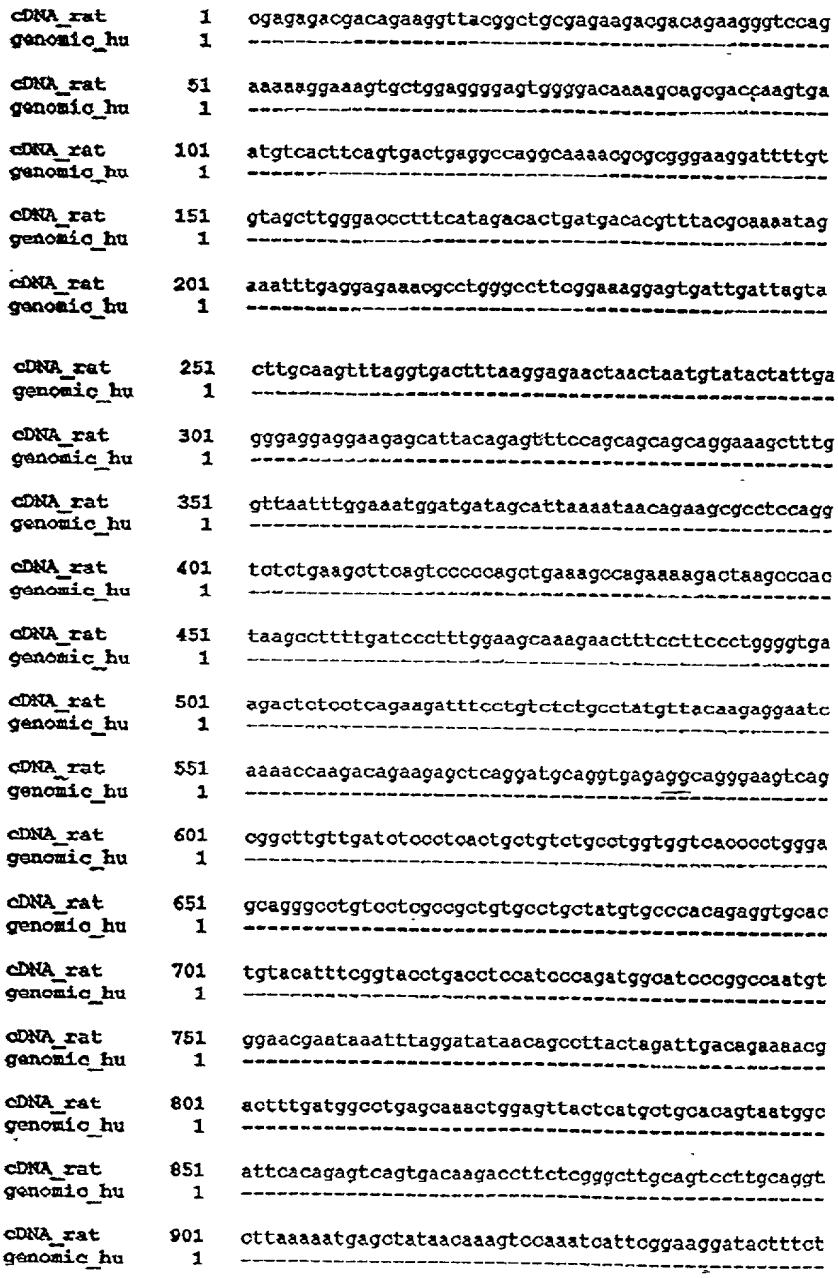


Figure 10


cDNA_rat 1 cgagagacgacagaaggttacggctgcgagaagacgacagaagggtccag
genomic_hu 1 ----------
cDNA_rat 51 aaaaaggaaagtgctggaggggagtggggacaaaagcagcgaccaagtga
genomic_hu 1 ----------
cDNA_rat 101 atgtcacttcagtgactgaggccaggcaaaacgcgcgggaaggattttgt
genomic_hu 1 ----------
cDNA_rat 151 gtagcttgggaccctttcatagacactgatgacacgtttacgcaaaatag
genomic_hu 1 ----------
cDNA_rat 201 aaatttgaggagaaacgcctgggccttcggaaaggagtgattgattagta
genomic_hu 1 ----------
cDNA_rat 251 cttgcaagtttaggtgactttaaggagaactaactaatgtatactattga
genomic_hu 1 ----------
cDNA_rat 301 gggaggaggaagagcattacagagtttccagcagcagcaggaaagctttg
genomic_hu 1 ----------
cDNA_rat 351 gttaatttggaaatggatgatagcattaaaataacagaagcgcctccagg
genomic_hu 1 ----------
cDNA_rat 401 tctctgaagcttcagtcccccagctgaaagccagaaaagactaagcccac
genomic_hu 1 ----------
cDNA_rat 451 taagcctttgatccctttggaagcaaagaactttccttccctggggtga
genomic_hu 1 ----------
cDNA_rat 501 agactctcctcagaagatttcctgtctctgcctatgttacaagaggaatc
genomic_hu 1 ----------
cDNA_rat 551 aaaaccaagacagaagagctcaggatgcaggtgagaggcagggaagtcag
genomic_hu 1 ----------
cDNA_rat 601 cggcttgttgatctccctcactgctgtctgcctggtggtcacccctggga
genomic_hu 1 ----------
cDNA_rat 651 gcagggcctgtcctcgccgctgtgcctgctatgtgcccacagaggtgcac
genomic_hu 1 ----------
cDNA_rat 701 tgtacatttcggtacctgacctccatcccagatggcatcccggccaatgt
genomic_hu 1 ----------
cDNA_rat 751 ggaacgaataaatttaggatataacagccttactagattgacagaaaacg
genomic_hu 1 ----------
cDNA_rat 801 actttgatggcctgagcaaactggagttactcatgctgcacagtaatggc
genomic_hu 1 ----------
cDNA_rat 851 attcacagagtcagtgacaagaccttctcgggcttgcagtccttgcaggt
genomic_hu 1 ----------
cDNA_rat 901 cttaaaaatgagctataacaaagtccaaatcattcggaaggatactttct
genomic_hu 1 ----------

```
cDNA_rat      951   acggactcgggagcttggtccggttgcacctggatcacaacaacattgaa
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1001   ttcatcaaccctgaggccttttatggacttacctcgctccgcttggtaca
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1051   tttagaaggaaaccggctcacaaagctccatccagacacatttgtctcat
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1101   taagctatctccagatatttaaaacctctttcattaagtacctgttcttg
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1151   tctgataacttcctgacctccctccccaaaagaaatggtctcctacatgcc
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1201   aaacctagaaagcctgtatttgcatggaaacccatggacctgtgactgcc
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1251   atttaaagtggttgtctgagtggatgcagggaaacccagatataataaaa
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1301   tgcaagaaagacagaagctcttccagtcctcagcaatgtccccctttgcat
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1351   gaaccccaggatctctaaaggcagacccctttgctatggtaccatctggag
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1401   ctttcctatgtacaaagccaaccattgatccatcactgaagtcaaagagc
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1451   ctggttactcaggaggacaatggatctgcctccacctcacctcaagattt
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1501   catagaaccctttggctccttgtctttgaacatgacananntntctggaa
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1551   ataaggccgacatggtctgtagtatccaaaagccatcaaggacatcacca
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1601   actgcattcactgaagaaaatgactacatcatgctaaatgcgtcattttc
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1651   cacaaatcttgtgtgcagtgtagattataatcacatccagccagtgtggc
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1701   aacttctggctttatacagtgactctcctctgatactagaaaggaagccc
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1751   cagcttaccgagactccttcactgtcttctagatataaacaggtggctct
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1801   taggcctgaagacatttttaccagcatagaggctgatgtcagagcagacc
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1851   ctttttggttccaacaagaaaaaattgtcttgcagctgaacagaactgcc
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1901   accacacttagcacattacagatccagttttccactgatgctcaaatcgc
genomic_hu      1   ----------------------------------------------------

cDNA_rat     1951   tttaccaagggcggagatgagagcggagagactcaaatggaccatgatcc
genomic_hu      1   ----------------------------------------------------
```

```
cDNA_rat      2001   tgatgatgaacaatcccaaactggaacgcactgtcctggttggcggcact
genomic_hu       1   --------------------------------------------------

cDNA_rat      2051   attgccctgagctgtccaggcaaaggcgacccttcacctcacttggaatg
genomic_hu       1   --------------------------------------------------

cDNA_rat      2101   gcttctagctgatgggagtaaagtgagagccccttacgttagcgaggatg
genomic_hu       1   --------------------------------------------------

cDNA_rat      2151   ggcgaatcctaatagacaaaaatgggaagttggaactgcagatggctgac
genomic_hu       1   --------------------------------------------------

cDNA_rat      2201   agctttgatgcaggtctttaccactgcataagcaccaatgatgcagatgc
genomic_hu       1   --------------------------------------------------

cDNA_rat      2251   ggatgttctcacatacaggataactgtggtagagccctatggagaaagca
genomic_hu       1   --------------------------------------------------

cDNA_rat      2301   cacatgacagtggagtccagcacacagtggttacgggtgagacgctcgac
genomic_hu       1   --------------------------------------------------

cDNA_rat      2351   cttccatgcctttccacgggtgttccagatgcttctattagctggattct
genomic_hu       1   --------------------------------------------------

cDNA_rat      2401   tccagggaacactgtgttctctcagccatcaagagacaggcaaattctta
genomic_hu       1   --------------------------------------------------

cDNA_rat      2451   acaatgggaccttaagaatattacaggttacgccaaaagatcaaggtcat
genomic_hu       1   --------------------------------------------------

cDNA_rat      2501   taccaatgtgtggctgccaacccatcaggggccgacttttccagttttaa
genomic_hu       1   --------------------------------------------------

cDNA_rat      2551   agtttcagttcaaaagaaaggccaaaggatggttgagcatgacagggagg
genomic_hu       1   --------------------------------------------------

cDNA_rat      2601   caggtggatctggacttggagaacccaactccagtgtttcccttaagcag
genomic_hu       1   --------------------------------------------------

cDNA_rat      2651   ccagcatctttgaaactctctgcatcagctttgacagggtcagaggctgg
genomic_hu       1   --------------------------------------------------

cDNA_rat      2701   aaaacaagtctccggtgtacataggaagaacaaacatagagacttaatac
genomic_hu       1   --------------------------------------------------

cDNA_rat      2751   atcggcggcgtggggattccacgctccggcgattcagggagcataggagg
genomic_hu       1   --------------------------------------------------

cDNA_rat      2801   cagctccctctctctgctcggagaattgacccgcaacgctgggcagcact
genomic_hu       1   --------------------------------------------------

cDNA_rat      2851   tctagaaaaagccaaaaagaattctgtgccaaaaaagcaagaaaatacca
genomic_hu       1   --------------------------------------------------

cDNA_rat      2901   cagtaaagccagtgccactggctgttcccctcgtggaactcactgacgag
genomic_hu       1   --------------------------------------------------

cDNA_rat      2951   gaaaaggatgcctctggcatgattcctccagatgaagaattcatggttct
genomic_hu       1   --------------------------------------------------
```

```
cDNA_rat      3001  gaaaactaaggcttctggtgtcccaggaaggtcaccaactgctgactctg
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3051  gaccagtaaatcatggttttatgacgagtatagcttctggcacagaagtc
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3101  tcaactgtgaatccacaaacactacaatctgagcaccttcctgatttcaa
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3151  attatttagtgtaacaaacggtacagctgtgacaaagagtatgaacccat
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3201  ccatagcaagcaaaatagaagatacaaccaaccaaaaacccaatcattatc
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3251  tttccatcagtagctgaaattcgagattctgctcaggcaggaagagcatc
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3301  ttcccaaagtgcacacctgtaacaggggaaacatggctacctatggcc
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3351  ataccaacacatatagtagctttaccagcaaagccagtacagtcttgcag
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3401  ccaataaatccaacagaaagttatggacctcagatacctattacaggagt
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3451  cagcagacctagcagtagtgacatctcttctcacactactgcagacccta
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3501  gcttctccagtcacccttcaggttcacacaccactgcctcgtctttattt
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3551  cacattcctagaaacaacaatacaggtaacttcccottgtccaggcactt
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3601  gggaagagagaggacaatttggagcagagggagagttaaaaacccacata
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3651  gaaccccagttctccgacggcatagacacaggactgtgaggccagcaatc
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3701  aagggacctgctaacaaaaatgtgagccaagttccagccacagagtacca
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3751  tgggatgtgccacacatgtccttccgcagaggggctcacagtggctactg
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3801  cagcactgtcagttccaagttcatcccacaagtgcctccccaaaactaat
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3851  aatgttggggtcatagcagaagagtctaccactgtggtcaagaaaccact
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3901  gttactatttaaggacaaacaaaatgtagatattgagataataacaacca
genomic_hu       1  ----------------------------------------------------

cDNA_rat      3951  ctacaaaatattccggaggggaaagtaaccacgtgattcctacggaagca
genomic_hu       1  ----------------------------------------------------

cDNA_rat      4001  agcatgacttctgctccaacatctgtatccctggggaaatctcctgtaga
genomic_hu       1  ----------------------------------------------------

cDNA_rat      4051  caatagtggtcacctgagcatgcctgggaccatccaaactgggaaagatt
genomic_hu       1  ----------------------------------------------------

cDNA_rat      4101  cagtggaaacaacaccacttcccagcccccctcagcacaccctcaatacca
genomic_hu       1  ----------------------------------------------------
```

```
cDNA_rat      4151  acaagcacaaaattctcaaagaggaaaactcccttgcaccagatctttgt
genomic_hu       1  ------------------------------------------------

cDNA_rat      4201  aaataaccagaagaaggagggatgttaaagaatccatatcaattcggtt
genomic_hu       1  ------------------------------------------------

cDNA_rat      4251  tacaaaagaacccagccgcaaagcttcccaaaatagctcctcttttaccc
genomic_hu       1  ------------------------------------------------

cDNA_rat      4301  acaggtcagagttcccccctcagattctacaactctcttgacaagtccgcc
genomic_hu       1  ------------------------------------------------

cDNA_rat      4351  accagctctgtctacaacaatggctgccactcagaacaagggcactgaag
genomic_hu       1  ------------------------------------------------

cDNA_rat      4401  tagtatcaggtgccagaagtctctcagcagggaagaagcagcccttcacc
genomic_hu       1  ------------------------------------------------

cDNA_rat      4451  aactcctctccagtgcttcctagcaccataagcaagagatctaatacatt
genomic_hu       1  ------------------------------------------------

cDNA_rat      4501  aaacttcttgtcaacggaaaccccccacagtgacaagtcctactgctactg
genomic_hu       1  ------------------------------------------------

cDNA_rat      4551  catctgtcattatgtctgaaacccaacgaacaagatccaaagaagcaaaa
genomic_hu       1  ------------------------------------------------

cDNA_rat      4601  gaccaaataaaggggcctcggaagaacagaaacaacgcaaacaccacccc
genomic_hu       1  ------------------------------------------------

cDNA_rat      4651  caggcaggtttctggctatagtgcatactcagctctaacaacagctgata
genomic_hu       1  ------------------------------------------------

cDNA_rat      4701  ccccttggctttcagtcattccccacgacaagatgatggtggaaatgta
genomic_hu       1  ------------------------------------------------

cDNA_rat      4751  agtgcagttgcttatcactcaacaacctctcttctggccataactgaact
genomic_hu       1  ------------------------------------------------

cDNA_rat      4801  gtttgagaagtacacccagactttgggaaatacaacagctttggaaacaa
genomic_hu       1  ------------------------------------------------

cDNA_rat      4851  cgttgttgagcaaatcacaggagagtaccacagtgaaaagagcctcagac
genomic_hu       1  ------------------------------------------------

cDNA_rat      4901  acaccaccaccactcctcagcagtggggcgcccccagtgcccactccttc
genomic_hu       1  ------------------------------------------------

cDNA_rat      4951  cccacctcctttactaagggtgtggttacagacagcaaagtcacatcag
genomic_hu       1  ------------------------------------------------

cDNA_rat      5001  ctttccagatgacgtcaaatagagtggtcaccatatatgaatcttcaagg
genomic_hu       1  ------------------------------------------------

cDNA_rat      5051  cacaatacagatctgcagcaaccctcagcagaggctagccccaatcctga
genomic_hu       1  ------------------------------------------------

cDNA_rat      5101  gatcataactggaaccactgactctccctctaatctgtttccatccactt
genomic_hu       1  ------------------------------------------------

cDNA_rat      5151  ctgtgccagcactaagggtagataaaccacagaattctaaatggaagccc
genomic_hu       1  ------------------------------------------------

cDNA_rat      5201  tctccctggccagaacacaaatatcagctcaagtcatactccgaaaccat
genomic_hu       1  ------------------------------------------------
```

```
cDNA_rat      5251   tgagaagggcaaaaggccagcagtaagcatgtccccccacctcagccttc
genomic_hu       1   ------------------------------------------------

cDNA_rat      5301   cagaggccagcactcatgcctcacactggaatacacagaagcatgcagaa
genomic_hu       1   ------------------------------------------------

cDNA_rat      5351   aagagtgtttttgataagaaacctggtcaaaaccccaacttccaaacatct
genomic_hu       1   ------------------------------------------------

cDNA_rat      5401   gccttacgtctctctacctaagactctattgaaaaagccaagaataattg
genomic_hu       1   ------------------------------------------------

cDNA_rat      5451   gaggaaaggctgcaagctttacagttccagctaattcagaogttttttctt
genomic_hu       1   ------------------------------------------------

cDNA_rat      5501   ccttgtgaggctgttggagacccactgcccatcatccactggaccagagt
genomic_hu       1   ------------------------------------------------

cDNA_rat      5551   ttcatcaggatttgaaatatcccaagggacacagaaaagccggttccacg
genomic_hu       1   ------------------------------------------------

cDNA_rat      5601   tgcttcccaatggcaccttgtccatccagagggtcagtattcaggaccgt
genomic_hu       1   ------------------------------------------------

cDNA_rat      5651   ggacagtacctgtgctctgcatttaatccactgggcgtagaccattttca
genomic_hu       1   ------------------------------------------------

cDNA_rat      5701   tgtctctttgtctgtggtttttacccggcaaggattttggacagacatg
genomic_hu       1   ------------------------------------------------

cDNA_rat      5751   tcaaggagatcacagttcactttggaagtactgtggaactaaagtgcaga
genomic_hu       1   ------------------------------------------------

cDNA_rat      5801   gtggagggtatgccgaggcctacggtttcctggatacttgcaaaccaaac
genomic_hu       1   ------------------------------------------------

cDNA_rat      5851   ggtggtctcagaaacggccaagggaagcagaaaggtctgggtaacacctg
genomic_hu       1   ------------------------------------------------

cDNA_rat      5901   atggaacattgatcatctataatctgagtctttatgatcgtggtttttac
genomic_hu       1   ------------------------------------------------

cDNA_rat      5951   aagtgtgtggccagcaacccatctggccaggattcactgttggttaagat
genomic_hu       1   ------------------------------------------------

cDNA_rat      6001   acaagtcatcacagctccccctgtcattatagagcaaaagaggcaagcca
genomic_hu       1   ------------------------------------------------

cDNA_rat      6051   tcgttggggttttaggtggaagtttgaaactgccctgcactgcaaaagga
genomic_hu       1   ------------------------------------------------

cDNA_rat      6101   actccccagcctagtgttcactgggtcctttatgatgggactgaactaaa
genomic_hu       1   ------------------------------------------------

cDNA_rat      6151   accattgcagttgactcattccagatttttcttgtatccaaatggaactc
genomic_hu       1   ------------------------------------------------

cDNA_rat      6201   tgtatataagaagcatcgctccttcagtgaggggcacttatgagtgcatt
genomic_hu       1   ------------------------------------------------

cDNA_rat      6251   gccaccagctcctcaggctcagagagaagggtagtgattcttactgtgga
genomic_hu       1   ------------------------------------------------
```

```
cDNA_rat      6301   agagggagagacaatccccaggatagaaactgcctctcagaaatggactg
genomic_hu       1   ------------------------------------------------------

cDNA_rat      6351   aggtgaatttgggtgagaaattactactgaactgctcagctactggggat
genomic_hu       1   ------------------------------------------------------

cDNA_rat      6401   ccaaagcctagaataatctggaggctgccatccaaggctgtcatcgacca
genomic_hu       1   ------------------------------------------------------

cDNA_rat      6451   gtggcacagaaTGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGg
genomic_hu       1   -----------TGGGCAGCTGGATCCACGTCTACCCTAATGGATCCCTGt

cDNA_rat      6501   tggttgggTCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCA
genomic_hu      40   ttattggaTCAGTAACAGAAAAAGACAGTGGTGTCTACTTGTGTGTGGCA

cDNA_rat      6551   AGAAACAAAATGGGAGATGACCTAGTCCTGATGCATGTccgcctgAGATT
genomic_hu      90   AGAAACAAAATGGGGGATGATCTGATACTGATGCATGTtagcctaAGACT

cDNA_rat      6601   GACACCTGCCAAAATTGAACAGAAGCAGTATTTTAagAAGCAAGTGCTCC
genomic_hu     140   GAAACCTGCCAAAATTGACCACAAGCAGTATTTTAgaAAGCAAGTGCTCC

cDNA_rat      6651   ATGGGAAAGATTTCCAAGTTGACTGCAaggcctctGGCTCCCCTGTGCCT
genomic_hu     190   ATGGGAAAGATTTCCAAGTAGATTGCAaagcttccGGCTCCCCAGTGCCA

cDNA_rat      6701   GAGGTATCCTGGAGTTTGCCTGATGggacagtgcTCAACAATGTAGCCCA
genomic_hu     240   GAGATATCTTGGAGTTTGCCTGATGgaaccatgaTCAACAATGCAATGCA

cDNA_rat      6751   AGCTGATGACAGTGGCTATAGGACcaagaggtacACCCTTTTCCACAATG
genomic_hu     290   AGCCGATGACAGTGGCCACAGGACtaggagatatACCCTTTTCAACAATG

cDNA_rat      6801   GAACCTTGTATTTCAACAACGTTGggatggcaGAGGAAGGAGATTATATC
genomic_hu     340   GAACTTTATACTTCAACAAAGTTGgggtagcgGAGGAAGGAGATTATACT

cDNA_rat      6851   TGCTCTGCCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAAC
genomic_hu     390   TGCTATGCCCAGAACACCCTAGGGAAAGATGAAATGAAGGTCCACTTAAC

cDNA_rat      6901   AGTTCTAACAGCcatcCCACGGATAAGGCAAAGCTACAAGACCACCATGA
genomic_hu     440   AGTTATAACAGCtgctCCCCGGATAAGGCAGAGTAACAAAACCAACAAGA

cDNA_rat      6951   GGCTCAggGCTGGAGAAACAGCTGTCCTTGACTGCGAGGTCACTGGGGaa
genomic_hu     490   GAATCAaaGCTGGAGACACAGCTGTCCTTGACTGTGAGGTCACTGGGGat

cDNA_rat      7001   ccgaagcccaatgTATTTTGGTTGCTGCCTTCCAAcaatgtcATTTCATT
genomic_hu     540   cccaaaccaaaaaTATTTTGGTTGCTGCCTTCCAATgacatgATTTCCTT

cDNA_rat      7051   CTCCAATGACAGGTTCACATTTCATGCCAATagaaCTTTGTCCATCCATA
genomic_hu     590   CTCCATTGATAGGTACACATTTCATGCCAATgggtCTTTGACCATCAACA

cDNA_rat      7101   AAGTGAAACCACTTGACTCTGGGGActatgtgtgcgtagctcagAATCCT
genomic_hu     640   AAGTGAAACTGCTCGATTCTGGAGAgtacgtatgtgtagcccgaAATCCC

cDNA_rat      7151   AGTGGGGATGACACTAAGACATACAAACTGGAcattGTCTCTAAACCTCC
genomic_hu     690   AGTGGGGATGACACCAAAATGTACAAACTGGAtgtgGTCTCTAAACCTCC

cDNA_rat      7201   ATTAATCAATGGCCTGTATGCAAACAagACTGTTATTAAAGCCACAGCca
genomic_hu     740   ATTAATCAATGGTCTGTATACAAACAgaACTGTTATTAAAGCCACAGCtg

cDNA_rat      7251   ttcggcacTCCAAAAAATACTTTGACTGCAGAGCAGATGGGATCCCATct
genomic_hu     790   tgagacatTCCAAAAAACACTTTGACTGCAGAGCTGAAGGGACACCATct

cDNA_rat      7301   tcccaggtcacGTGGATTATGCCAGGCAATATTTTCCTCCCAGCTCCATA
genomic_hu     840   cctgaagtcatGTGGATCATGCCAGACAATATTTTCCTCACAGCCCCATA
```

cDNA_rat	7351	CttTGGAAGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAgatgagga
genomic_hu	890	CtaTGGAAGCAGAATCACAGTCCATAAAAATGGAACCTTGGAaattagga
cDNA_rat	7401	acatccGGCTTTCTGACTCTGCGGACTTCACCTGTGTGGttcggagcgag
genomic_hu	940	atgtgaGGCTTTCAGATTCAGCCGACTTTATCTGTGTGGcccgaaatgaa
cDNA_rat	7451	ggaGGAGAGAGTGTGTTGGTAGTGCAGTTAGAAGTCCTAGAAATGCTGAG
genomic_hu	990	ggtGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAG
cDNA_rat	7501	AAGACCAACATTCAGAAACCCATTCAACGAAAAagtcatcgcccaagctg
genomic_hu	1040	AAGACCGACATTTAGAAATCCATTTAATGAAAAaatagttgcccagctgg
cDNA_rat	7551	gcaagcccgtAGCACTGAACTGCTCTGTGGATGGGAACCCCCCACCTGAA
genomic_hu	1090	gaaagtccacAGCATTGAATTGCTCTGTTGATGGTAACCCACCACCTGAA
cDNA_rat	7601	attacCTGGATCTTACCTGACGGCACACAGTTTGCTAACAGACCAcacaa
genomic_hu	1140	ataatCTGGATTTTACCAAATGGCACACGATTTTCCAATGGACCAcaaag
cDNA_rat	7651	ttccccGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAAgcaa
genomic_hu	1190	ttatcaGTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCTAAAacaa
cDNA_rat	7701	ctcggaacaagtCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGC
genomic_hu	1240	ctcgggaggatgCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGC
cDNA_rat	7751	TAcatcgagaaactcatcctgttagagATTGGGCAGAAGCCAGTCATTCT
genomic_hu	1290	TAtattgagaaattagtcatattagaaATTGGCCAGAAGCCAGTTATTCT
cDNA_rat	7801	GACATACGAACCAGGGAtggtgaagagcgTCAGTGGGGAACCGTTATCAC
genomic_hu	1340	TACCTATGCACCAGGGAcagtaaaaggcaTCAGTGGAGAATCTCTATCAC
cDNA_rat	7851	TGCATTGTGTGTCTGATGGgATCCCCAAGCCAAATGTCAAGTGGACTACA
genomic_hu	1390	TGCATTGTGTGTCTGATGGaATCCCTAAGCCAAATATCAAATGGACTATG
cDNA_rat	7901	CCGGGTGGccATGTAATCGACAGGCCTCAAGTGGATGGAAAATACATAcT
genomic_hu	1440	CCAAGTGGttATGTAGTAGACAGGCCTCAAATTAATGGGAAATACATAtT
cDNA_rat	7951	GCATGAAAATGGCACGCTGGTCATCAAAGCAACAACAGCtcacgaccaAG
genomic_hu	1490	GCATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCttatgacagAG
cDNA_rat	8001	GAAATTATATCTGTAGGGCTCAAAACAGTGTTGGCCAggcagttattagc
genomic_hu	1540	GAAACTATATCTGTAAGGCTCAAAATAGTGTTGGTCAtacactgattact
cDNA_rat	8051	gtgtCAGTGATGGTTGTGGCCTACCCTCCCCGAATCATAAActacctACC
genomic_hu	1590	gttcCAGTAATGATTGTAGCCTACCCTCCCCGAATTACAAAtcgtccACC
cDNA_rat	8101	CAGGAACATGCTCAGGAGGACAGGGGAAGCCatgCAGCTCCACTGTGTGG
genomic_hu	1640	CAGGAGTATTGTCACCAGGACAGGGGCAGCCtttCAGCTCCACTGTGTGG
cDNA_rat	8151	CCTTGGGAATCCCCAAGCCAaAAGTCACCTGGGAGACGCCAAGACACTCC
genomic_hu	1690	CCTTGGGAGTTCCCAAGCCAgAAATCACATGGGAGATGCCTGACCACTCC
cDNA_rat	8201	CTGCTCTCAAaagcaacagcaagaaaacccCATAGAAGTGAGATGCTTCA
genomic_hu	1740	CTTCTCTCAAcggcaagtaaagagaggacaCATGGAAGTGAGCAGCTTCA
cDNA_rat	8251	CCCACAAGGTACGCTgGTCATTCAGAATCTCCAAACCTCGGATTCCGGag
genomic_hu	1790	CTTACAAGGTACCCTaGTCATTCAGAATCCCCAAACCTCCGATTCTGGga
cDNA_rat	8301	tcTATAAGTGCAGAGCTCAGAACCTACTTGGgacTGATTACGCAACAACT
genomic_hu	1840	taTACAAATGCACAGCAAAGAACCCACTTGGtagTGATTATGCAGCAACG
cDNA_rat	8351	TACATCCAGGTACTCTGACAGGAAgggggagactaaaattcaacagaagt
genomic_hu	1890	TATATTCAAGTAATCTGACATGAAataataaagtcaacaacatctgggca

```
cDNA_rat      8401   ccacatccacagggTTTATTTTTTGGAAGAAGTTTAATCAAAGGCAGCCA
genomic_hu    1940   gaa----------TTTATTTTTTGGAAGAAGTTTAATCAAAGGCAGCCA

cDNA_rat      8451   TAGGCATGTAAATGAgtcTGAATACATTTACAGTATTAAATTTACAATGG
genomic_hu    1979   TAGGCATGTAAATGAattTGAATACATTTACAGTATTAAATTTACAATGA

cDNA_rat      8501   ACATGCgatga-----GACTTGTAAATGAAAGCATTGTGAACTGAaaccg
genomic_hu    2029   ACATGCaaaataaaagGACTTGTAAATAAATGCATTATGAACTGAtgata

cDNA_rat      8546   agtctctg-----TGGATCTCAAAGCAAACTCTTAACTTAAGGCACTTTg
genomic_hu    2079   ctgatttatttaaTGGATCTCAAAACAAACTTTTAACTTAAGGCACTTTt

cDNA_rat      8591   ATTTTGCCAACAAATAATAACAAACAttaagagaaaaaaatgatcCACTA
genomic_hu    2129   ATTTTGCCAACAAATAACAATAAACAaacattgaaacggtt----CACTA

cDNA_rat      8641   CGAAATAACAAACGGCTAATGCACCTGAATTctcagtaaaaagacctttc
genomic_hu    2175   TAAAATAACAAATGGCTAATGTACCTGAATTtttcagtaaaaaaatgaac

cDNA_rat      8691   tctcgctaacagttgCCAGCTGCCTCGTGTCTGTTTCCTACCAATGTCAC
genomic_hu    2225   ttctaata-------CCAGTTGCCTAGTGTCCACCTCCTATCAATGTTAC

cDNA_rat      8741   AAACAtcgcacacagggtGAATGGAGTCAACGGGAAAGATTAAGTTTGCG
genomic_hu    2268   AAGCAtggcactca----GAACAGAGACAATGGAAAATATTAAATCTGCA

cDNA_rat      8791   GTCtgtgtaaatctca------------ATGTACAAATATTCTGtcncTG
genomic_hu    2314   ATCtttatgatgtaaatttaccatcctgATGTATAAATATTTTG----TG
```

exon1 (2342-2397)

```
cDNA_rat      8791   GTCtgtgtaaatctca------------ATGTACAAATATTCTGtcncTG
genomic_hu    2314   ATCtttatgatgtaaatttaccatcctgATGTATAAATATTTTG----TG

cDNA_rat      8829   GTTTATAAACATTTTGATAAAACCGAAAAAAAAAAAAAaaaaaaaaaaaa
genomic_hu    2360   GTTTATAAATTTTTTTGCTAAAACCTACAGAAAATAAGcactgaactgtc
```

(Genomic human OCP: SEQ ID NO:4)
(cDNA rat: SEQ ID NO:5)

Figure 11

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length	Remarks
1	1	208	208	69	No valid splice site found upstream this exon
2	277	429	153	18	
3	447	485	39	1561	
4	2046	2244	199	1351	
5	3595	3724	130	3254	
6	6978	7359	382	4123	
7	11482	14903	3422	38	
8	14941	15307	367	51	
9	15358	15825	468	1039	
10	16864	~17760	~897	?	Last exon might be complete
11	?(1)	2317	~2317	25	Exon is not complete and start site is not known
12	2342	2397	56	-	

FIGURE 12

608 Human translated nucleotide sequence (ORF)

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGT
GATCTGCCTGGTCGCCACCCCTGGGGGCAAGGCCTGTCCTCGCCGCTGTG
CCTGTTATATGCCTACGGAGGTACACTGCACATTTCGGTACCTGACTTCCA
TCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAAC
AGCTTGGTTAGATTGATGGAAACAGATTTTTCTGGCCTGACCAAACTGGA
GTTACTCATGCTTCACAGCAATGGCATTCACACAATCCCTGACAAGACCTT
CTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTATAATAAAGTCC
GAAAACTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGATTG
CACATGGACCACAACAATATTGAGTTTATAAACCCAGAGGTTTTTTATGG
GCTCAACTTTCTCCGCCTGGTGCACTTGGAAGGAAATCAGCTCACTAAGCT
CCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAAATCTC
TTTCATTAAGTTCCTATACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAA
GAGATGGTCTCCTATATGCCTGACCTAGACAGCCTTTACCTGCATGGAAA
CCCATGGACCTGTGATTGCCATTTAAAGTGGTTGTCTGACTGGATACAGCC
AGATGTAATAAAATGCAAAAAAGATAGAAGTCCCTCTAGTGCTCAGCAGT
GTCCACTTTGCATGAACCCTAGGACTTCTAAAGGCAAGCCGTTAGCTATG
GTCTCAGCTGCAGCTTTCCAGTGTGCCAAGCCAACCATTGACTCATCCCTG
AAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTTCATCTCT
CCCCAAGGTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGAT
CAGTCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAG
GACATCACCCATTGCATTCACTGAAGAAAATGACTACATCGTGCTAAATA
CTTCATTTTCAACATTTTTGGTGTGCAACATAGATTACGGTCACATTCAGC
CAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATACTAGAAA
GGAGCCACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAG
GTGGCTCCTAAGCCTGAAGACATTTTTACCAACATAGAGGCAGATCTCAG
AGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCCTTGCAGCTGAACA
GAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCT
CAAATCACTTTACCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGA
CTATGATTTCAAGGGATAACAATACTAAGCTGGAACATACTGTCTTGGTA
GGTGGAACCGTTGGCCTGAACTGCCCAGGCCAAGGAGACCCCACCCCACA
CGTGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCTTATGTCA
GTGAGGATGGACGGATCCTAATAGACAAAAGTGGAAAATTGGAACTCCA
GATGGCTGATAGTTTTGACACAGGCGTATATCACTGTATAAGCAGCAATT
ATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGG
TCGAAGCCTATCAGGAAAATGGGATTCATCACACAGTTTTCATTGGTGAA
ACACTTGATCTTCCATGCCATTCTACTGGTATCCCAGATGCCTCTATTAGC
TGGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAA
AGTTCTAAACAATGGCACATTAAGAATATTACAGGTCACCCCGAAAGACC
AAGGTTATTATCGCTGTGTGGCAGCCAACCCATCAGGGGTTGATTTTTGA
TTTTCCAAGTTTCAGTCAAGATGAAAGGACAAAGGCCCTTGGAGCATGAT
GGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCT
TAAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG
AGGTTGGAAAACACACCTCAAGCACAAGTAAGAGGCACAACTATCGGGA
ATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGA
ATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGG
GCGGCACTGTTGGAGAAAGCTAAAAAGAATGCTATGCCAGACAAGCGAG
AAAATACCACAGTGAGCCCACCCCCAGTGGTCACCCAACTCCCAAACATA
CCTGGTGAAGAAGACGATTCCTCAGGCATGCTCGCTCTACATGAGGAATT

TATGGTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTG
CTGACTCCAGAACAATATCTGATAGTCCTATGACAAACATAAATTATGGC
ACAGAATTCTCTCCTGTTGTGAATTCACAAATACTACCACCTGAAGAACCC
ACAGATTTCAAACTGTCTACTGCTATTAAAACTACAGCCATGTCAAAGAA
TATAAACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATT
CATCCACTGTCTTTCCACTGCTACTTGGAGCAACTGAATTTCAGGACTCTG
ACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCCCCAATAAC
AGTAAGGACTATGATCAAAGATGTCAATGTCAAAATGCTTAGTAGCACCA
CCAACAAACTATTATTAGAGTCAGTAAATACCACAAATAGTCATCAGACA
TCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCACTTCTATTCTCACAC
TACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCTGC
TCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTT
CAGACGCTTTGGGAGGCAGAGGAAAATTGGCGGAAGGGGGCGGATTATC
AGCCCATATAGAACTCCAGTTCTGCGACGGCATAGATACAGCATTTTCAG
GTCAACAACCAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCA
CAGTGCTCAATGTGACATGTCTGTCCTGTCTTCCCAGGGAGAGGCTCACCA
CTGCCACAGCAGCATTGTCTTTTCCAAGTGCTGCTCCCATCACCTTCCCCA
AAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACTCTAGTCCAG
AATCCACTATTACTACTTGAGAACAAACCCAGTGTAGAGAAAACAACACC
CACAATAAAATATTTCAGGACTGAAATTTCCCAAGTGACTCCAACTGGTG
CAGTCATGACATATGCTCCAACATCCATACCCATGGAAAAAACTCACAAA
GTAAACGCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGA
TTCAGTGATTACATCGTCACTTTCAGGTGCTATCACCAAGCCACCAATGAC
TATTATAGCCATTACAAGGTTTTCAAGAAGGAAAATTCCCTGGCAACAGA
ACTTTGTAAATAACCATAACCCAAAAGGCAGATTAAGGAATCAACATAAA
GTTAGTTTACAAAAAAGCACAGCTGTGATGCTTCCTAAAACATCTCCTGCT
TTACCACAGAGACAAAGTTCCCCTTTCCATTTCACCACACTTTCAACAAGT
GTGATGCAAATTCCATCTAATACCTTGACTACCGCTCACCACACTACGACC
AAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCCTTCCC
ACCCCTTAACCCTATGCTTCCTAGTATTATAAGCAAAGACTCAAGTACAA
AAAGCATCATATCAACGCAAACAGCAATACCAGCAACAACTCCTACCTTC
CCTGCATCTGTCATCACTTATGAAACCCAAACAGAGAGATCTAGAGCACA
AACAATACAAAGAGAACAGGAGCCTCAAAAGAAGAACAGGACTGACCCA
AACATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACA
CCTCCTGCTCTGGCATTCACTCATTCCCCACCAGAAAACACAACTGGGATT
TCAAGCACAATCAGTTTTCATTCAAGAACTCTTAATCTGACAGATGTGATT
GAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTC
TGAAACAACTTTGTCCAGCAAATCACACCAGAGTACCACAACTAGGAAAG
CATCATTAGACACTCCCATACCACCATTCTTGAGCAGCAGTGCTACTCTAA
TGCCAGTTCCCATCTCCCCTCCCTTTACTCAGAGAGCAGTTACTGACACAC
GTGGCGACTCCCATTTCCGGCTTATGACAAATACAGTGGTCAAGCTGCAC
GAATCCTCAAGGCACAATCTCCAAATGCCAAGTTCACAATTGGAACCACT
CACTTCATCTACCTCTAATCTGTTACATTCTACTCCCATGCCAGCACTAAC
AACAGTTAAATCACAGAATTCCAAATTAACTCCATCTCCCTGGGCAGAAT
ACCAATTTTGGCACAAACCATACTCAGACATTGCTGAAAAAGGCAAAAAG
CCAGAAGTAAGCATGTTGGCTACTACAGGCCTGTCCGAGGCCACCACTCT
TGTTTCAGATTGGGATGGACAGAAGAACACAAAGAAGAGTGACTTTGATA
AGAAACCAGTTCAAGAAGCAACAACTTCCAAACTCCTTCCCTTTGACTCTT
TGTCTAGGTATATATTTGAAAAGCCCAGGATAGTTGGAGGAAAAGCTGCA
AGTTTTACTATTCCAGCTAACTCAGATGCCTTTCTTCCCTGTGAAGCTGTT

GGAAATCCCCTGCCCACCATTCATTGGACCAGAGTTTCAGGACTTGATTTA
TCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCCAATGGTACCCT
GTCCATCCAGAGGGTGGAAATTCAGGACCGCGGACAGTACTTGTGTTCCG
CATCCAATCTGTTTGGCACAGACCACCTTCATGTCACCTTGTCTGTGGTTT
CCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTTCAT
TCCGGAAGCACTGTGGAACTGAAGTGCAGAGCAGAAGGTAGGCCAAGCC
CTACAGTTACCTGGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCC
AGGGAAGTAGGCAGGCTGTGGTGACGGTTGACGGAACATTGGTCCTCCAC
AATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAACCCA
GGTGGCCAGGATTCACTGCTGGTTAAAATACAAGTCATTGCAGCACCACC
TGTTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGCACTTGGGGTGAAA
GTTTAAAACTGCCCTGTACTGCAAAAGGAACTCCTCAGCCCAGCGTTTACT
GGGTCCTCTCTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCCA
AGTTGTTCTTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTT
CAGACAGGGGCACTTATGAATGCATTGCTACCAGTTCCACTGGTTCGGAG
CGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACCAGCCCCAGGA
TAGAAGCTGCATCCCAGAAAAGGACTGAAGTGAATTTTGGGGACAAATTA
CTACTGAACTGCTCAGCCACTGGGGAGCCCAAACCCCAAATAATGTGGAG
GTTACCATCCAAGGCTGTGGTCGACCAGTGGAGCTGGATCCACGTCTACC
CTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTC
TACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATGCA
TGTTAGCCTAAGACTGAAACCTGCCAAAATTGACCACAAGCAGTATTTTA
GAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGCAAAGCTTCC
GGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGAT
CAACAATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATAT
ACCCTTTTCAACAATGGAACTTTATACTTCAACAAAGTTGGGGTAGCGGA
GGAAGGAGATTATACTTGCTATGCCCAGAACACCCTAGGGAAAGATGAA
ATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATAAGGCAGAG
TAACAAAACCAACAAGAGAATCAAAGCTGGAGACACAGCTGTCCTTGACT
GTGAGGTCACTGGGGATCCCAAACCAAAAATATTTTGGTTGCTGCCTTCC
AATGACATGATTTCCTTCTCCATTGATAGGTACACATTTCATGCCAATGGG
TCTTTGACCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATG
TGTAGCCCGAAATCCCAGTGGGGATGACACCAAAATGTACAAACTGGATG
TGGTCTCTAAACCTCCATTAATCAATGGTCTGTATACAAACAGAACTGTTA
TTAAAGCCACAGCTGTGAGACATTCCAAAAAACACTTTGACTGCAGAGCT
GAAGGGACACCATCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTT
CCTCACAGCCCCATACTATGGAAGCAGAATCACAGTCCATAAAAATGGAA
CCTTGGAAATTAGGAATGTGAGGCTTTCAGATTCAGCCGACTTTATCTGTG
TGGCCCGAAATGAAGGTGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTA
CTGGAAATGCTGAGAAGACCGACATTTAGAAATCCATTTAATGAAAAAAT
AGTTGCCCAGCTGGGAAAGTCCACAGCATTGAATTGCTCTGTTGATGGTA
ACCCACCACCTGAAATAATCTGGATTTTACCAAATGGCACACGATTTTCCA
ATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCA
TTTCTAAAACAACTCGGGAGGATGCAGGAAAATATCGCTGTGCAGCTAGG
AATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAA
GCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAG
AATCTCTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCA
AATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCCTCAAATTAATGGG
AAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGC
TTATGACAGAGGAAACTATATCTGTAAGGCTCAAAATAGTGTTGGTCATA

CACTGATTACTGTTCCAGTAATGATTGTAGCCTACCCTCCCCGAATTACAA
ATCGTCCACCCAGGAGTATTGTCACCAGGACAGGGGCAGCCTTTCAGCTC
CACTGTGTGGCCTTGGGAGTTCCCAAGCCAGAAATCACATGGGAGATGCC
TGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGAAGTG
AGCAGCTTCACTTACAAGGTACCCTAGTCATTCAGAATCCCCAAACCTCC
GATTCTGGGATATACAAATGCACAGCAAAGAACCCACTTGGTAGTGATTA
TGCAGCAACGTATATTCAAGTAATCTGA

Figure 13

A

Region						
General	Rat	Human	Region Length	% identity	% positives	% gaps
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region					
Rat	Mouse	Length	% identity	% positives	% gaps
1-238	1-238	238	91	92	1

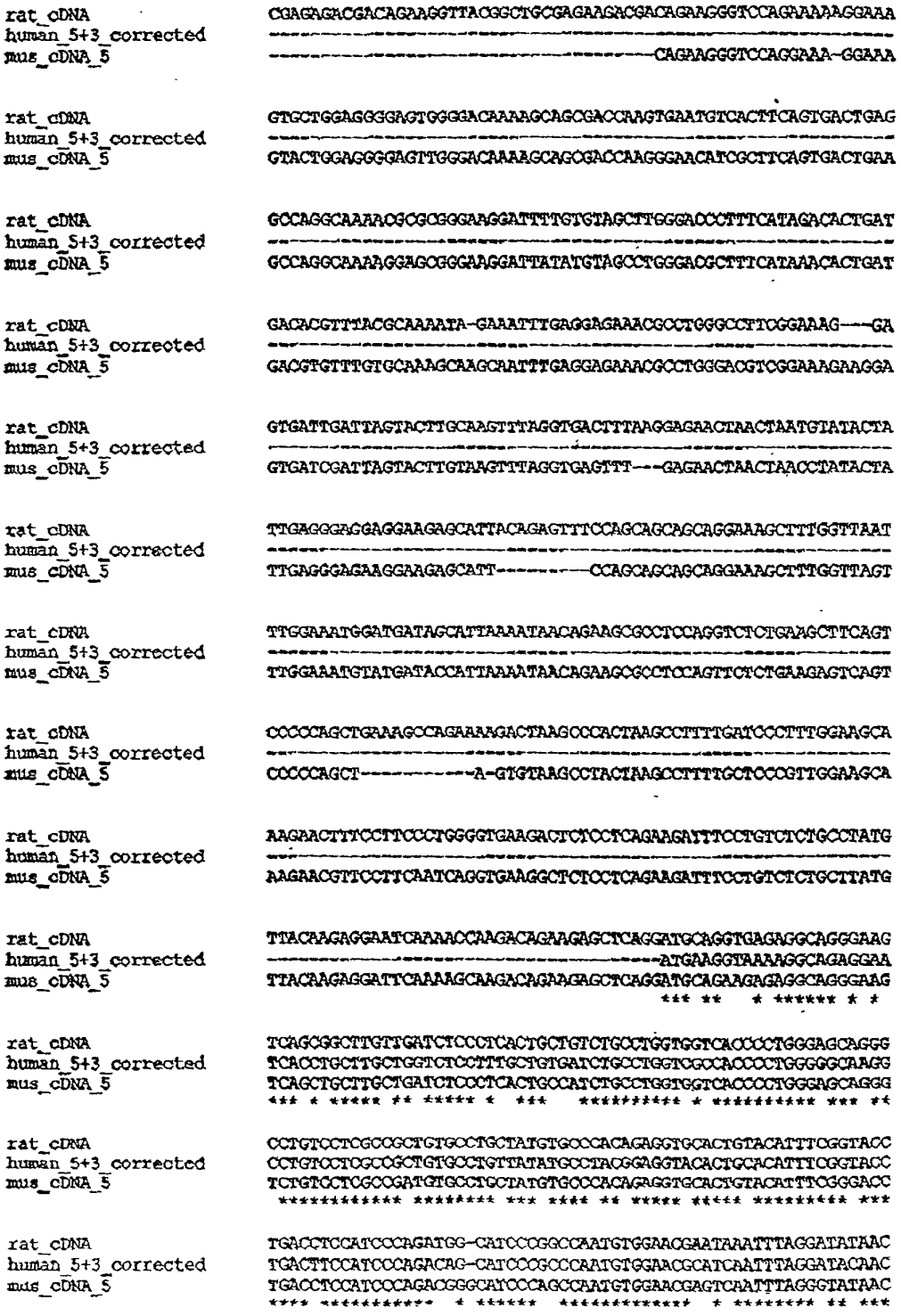
C

Region				
General	Rat	Human	Region Length	% identity
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

D

Region				
General	Rat	Mouse	Region Length	% identity
1-720	1-718	1-720	720	93

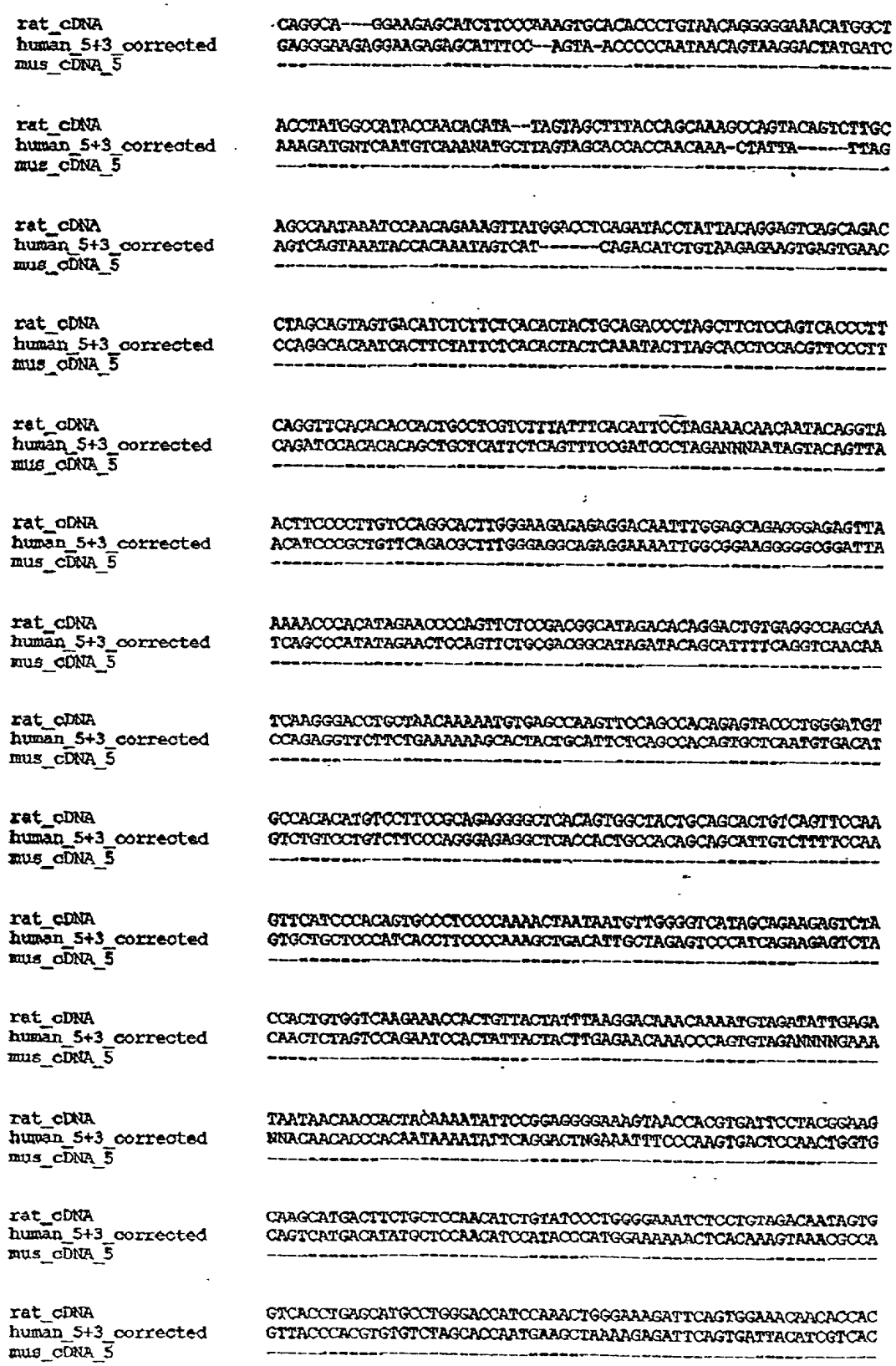
Figure 14


rat_cDNA
human_5+3_corrected
mus_cDNA_5
CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAAGGAAA
----------CAGAAGGGTCCAGGAAA-GGAAA
GTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACTTCAGTGACTGAG
GTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAA
GCCAGGCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTTTCATAGACACTGAT
GCCAGGCAAAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGAT
GACACGTTTACGCAAAATA-GAAATTTGAGGAGAAACGCCTGGGCCTTCGGAAAG---GA
GACGTGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAGGA
GTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAACTAACTAATGTATACTA
GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTT---GAGAACTAACTAACCTATACTA
TTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGCAGCAGGAAAGCTTTGGTTAAT
TTGAGGGAGAAGGAAGAGCATT---------CCAGCAGCAGCAGGAAAGCTTTGGTTAGT
TTGGAAATGGATGATAGCATTAAAATAACAGAAGCGCCTCCAGGTCTCTGAAGCTTCAGT
TTGGAAATGTATGATACCATTAAAATAACAGAAGCGCCTCCAGTTCTCTGAAGAGTCAGT
CCCCCAGCTGAAAGCCAGAAAAGACTAAGCCCACTAAGCCTTTTGATCCCTTTGGAAGCA
CCCCCAGCT----------A-GTGTAAGCCTACTAAGCCTTTTGCTCCCGTTGGAAGCA
AAGAACTTTCCTTCCCTGGGGTGAAGACTCTCCTCAGAAGATTTCCTGTCTCTGCCTATG
AAGAACGTTCCTTCAATCAGGTGAAGGCTCTCCTCAGAAGATTTCCTGTCTCTGCTTATG
TTACAAGAGGAATCAAAACCAAGACAGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAG
----------ATGAAGGTAAAAGGCAGAGGAA
TTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAGGATGCAGAAGAGAGGCAGGGAAG
TCAGCGGCTTGTTGATCTCCCTCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGG
TCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACCCCTGGGGGCAAGG
TCAGCTGCTTGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGG
CCTGTCCTCGCCGCTGTGCCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCGGTACC
CCTGTCCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTCGGTACC
TCTGTCCTCGCCGATGTGCCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCGGGACC
TGACCTCCATCCCAGATGG-CATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAAC
TGACTTCCATCCCAGACAG-CATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAAC
TGACCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTATAAC

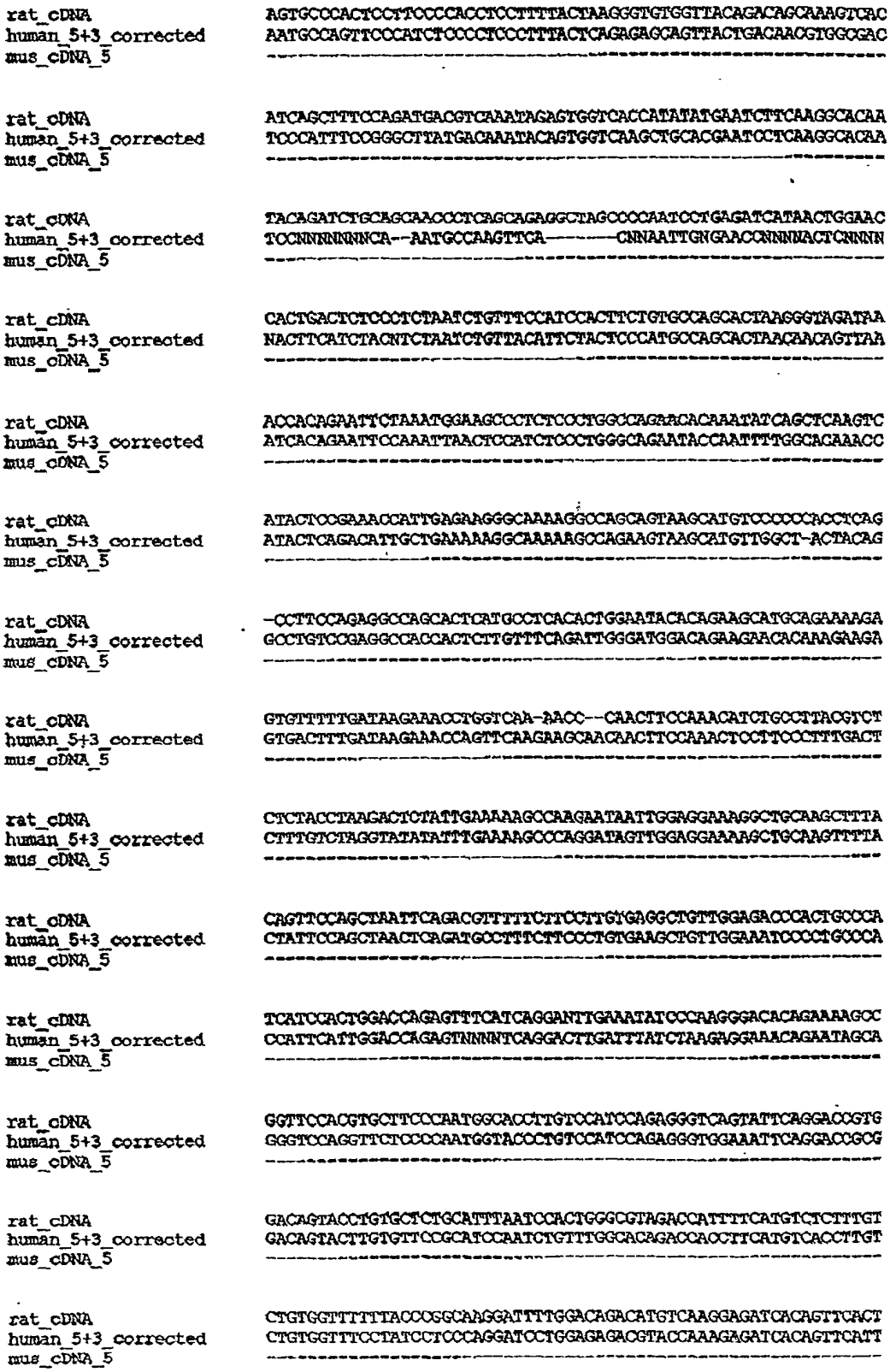
```
rat_cDNA                AGCCTTACTAGATTGACAGAAAACGACTTTGATGGCCTGAGCAAACTGGAGTTACTCATG
human_5+3_corrected     AGCTTGGTTAGATTGATGGAAACAGATTTTTCTGGCCTGACCAAACTGGAGTTACTCATG
mus_cDNA_5              AGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCTGAGCAGACTGGAGTTACTCATG
                        *** *   ********  ****  ** ***  ******** ** ******************

rat_cDNA                CTGCACAGTAATGGCATTCACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAGTCCTTG
human_5+3_corrected     CTTCACAGCAATGGCATTCACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTG
mus_cDNA_5              CTGCACAGCAATGGCATTCACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAGTCCTTG
                        ** ***** ************** * **  ***************  *  ******* *****

rat_cDNA                CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCATTCGGAAGGATACTTTCTACGGA
human_5+3_corrected     CAGGTCTTAAAAATGAGCTATAATAAAGTCCGAAAACTTCAGAAAGATACTTTTTATGGC
mus_cDNA_5              CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATAATTGAGAAGGATACTTTGTATGGA
                        ************************ ******* **    **  *** ******** ** **

rat_cDNA                CTCGGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG
human_5+3_corrected     CTCAGGAGCTTGACACGATTGCACATGGACCACAACAATATTGAGTTTATAAACCCAGAG
mus_cDNA_5              CTCAGGAGCTTGACCCGGTTGCACCTGGATCACAACAACATTGAGTTTATCAACCCCGAG
                        *** ********    ** ****** **** ******** ***** ** ** ***** ***

rat_cDNA                GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACAAAG
human_5+3_corrected     GTTTTTTATGGGCTCAACTTTCTCCGCCTGGTGCACTTGGAAGGAAATCAGCTCACTAAG
mus_cDNA_5              GCGTTTTACGGACTCACCTTGCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAG
                        *  ***** ** ** * **  ****** **** **  * ******** * *** ** ***

rat_cDNA                CTCCATCCAGACACATTTGTCTCATTAAGCTATCTCCAGATATTTAAAACCTCTTTCATT
human_5+3_corrected     CTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAAATCTCTTTCATT
mus_cDNA_5              CTCCATCCAGACACATTTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATT
                        ***** ***** ************ ** ****** ***************** *** ******

rat_cDNA                AAGTACCTGTTCTTGTCTGATAACTTCCT-GACCTCCCTCCCAAAAGAAATGGTCTCCTA
human_5+3_corrected     AAGTTCCTATACTTGTCTGATAACTTCCT-GACCTCCCTCCCTCAAGAGATGGTCTCCTA
mus_cDNA_5              AAGNACCTGTACTTGTATGATAACTTCATTGACCTCCCTCCCAAAAGAAATGGTCTCCTC
                        ***  *** * ***** ********** * ************  **** **********

rat_cDNA                CATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGACTGCCATTT
human_5+3_corrected     TATGCCTGACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTT
mus_cDNA_5              TATGCCAAACCTAGAAAGCCTTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATTT
                         *****  ******* *****  **  ************************** ********

rat_cDNA                AAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAATGCAAGAAAGACAG
human_5+3_corrected     AAAGTGGTTGTCTGACTGGATACAGGNNNNNCCAGATGTAATAAAATGCAAAAAAGATAG
mus_cDNA_5              AAAGTGGTTGTCCGAGTGGATGCAGGGAAACCCAGGTA-ACTATCTTGT-----TTGTTTG
                        ************ ** ***** ****     **** *  * **   **       *  *

rat_cDNA                AAGCTCTTCCAGTCCTCAGCAATGTCCCCT-TTGCATGAACCCCAGGATCTCTAAAGGCA
human_5+3_corrected     AAGTCCCTCTAGTGCTCAGCAGTGTCCACT-TTGCATGAACCCTAGGACTTCTAAAGGCA
mus_cDNA_5              TTTCTTTTTTTATARKACGTATTTTCCTCAATTTCATTTAGAATGATATCCCAAAAGTC-
                              *    *      * * * *** *  ** ***  *       *   * **** *

rat_cDNA                GACCCTTTGCTATGGTACCATCTGGAGCTTTCCTATGTACAAAGCCAACCATTGATCCAT
human_5+3_corrected     AGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAGTGTGCCAAGCCAACCATTGACTCAT
mus_cDNA_5              -CCCCATAACCTCCCCCCCA--------CTTCCCTACCTACCCATTC--CCATTTTTTGGC
                          **  *  *          **        *** **    * *  *  *  *****

rat_cDNA                CACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAATGGATCTGCCTCCACCTCACCTC
human_5+3_corrected     CCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTTCATCTCTCCCC
mus_cDNA_5              CCTGGCATTCCCC----------------------------------------------
                        *   * * **

rat_cDNA                AAGATTTCATAGAACCCTTTGGCTCCTTGTCTTTGAACATGACANANNTNTCTGGAAATA
human_5+3_corrected     AAGGTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGATCAGTCTGGAAATG
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                AGGCCGACATGGTCTGTAGTATCCAAAAGCCATCAAGGACATCACCAACTGCATTCACTG
human_5+3_corrected     AAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAGGACATCACCCATTGCATTCACTG
mus_cDNA_5              ----------------------------------------------------------
```

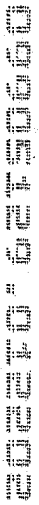

rat_cDNA AAGAAAATGACTACATCATGCTAAATGCGTCATTTTCCACAAATCTTGTGTGCAGTGTAG
human_5+3_corrected AAGAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTTGGTGTGCAACATAG
mus_cDNA_5
rat_cDNA ATTATAATCACATCCAGCCAGTGTGGCAACTTCTGGCTTTATACAGTGACTCTCCTCTGA
human_5+3_corrected ATTACGGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGA
mus_cDNA_5
rat_cDNA TACTAGAAAGGAAGCCCCAGCTTACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGG
human_5+3_corrected TACTAGAAAGGAGCCACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGG
mus_cDNA_5
rat_cDNA TGGCTCTTAGGCCTGAAGACATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTT
human_5+3_corrected TGGCTCCTAAGCCTGAAGACATTTTTACCAACATAGAGGCAGATCTCAGAGCAGATCCCT
mus_cDNA_5
rat_cDNA TTTGGTTCCAACAAGAAAAAATTGTCTTGCAGCTGAACAGAACTGCCACCACACTTAGCA
human_5+3_corrected CTTGGTTAATGCAAGACCAAATTTCCTTGCAGCTGAACAGAACTGCCACCACATTCAGTA
mus_cDNA_5
rat_cDNA CATTACAGATCCAGTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAG
human_5+3_corrected CATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTACCAAGAGCAGAGATGAGGC
mus_cDNA_5
rat_cDNA CGGAGAGACTCAAATGGACCATGATCCTGATGATGAACAATCCCAAACTGGAACGCACTG
human_5+3_corrected CAGTGAAACACAAATGGACTATGATTCAAGGGATAACAATACTAAGCTGGAACATACTG
mus_cDNA_5
rat_cDNA TCCTGGTTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTCACCTCACT
human_5+3_corrected TCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCCAGGCCAAGGAGACCCCACCCCACACG
mus_cDNA_5
rat_cDNA TGGAATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGGC
human_5+3_corrected TGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCTTATGTCAGTGAGGATGGAC
mus_cDNA_5
rat_cDNA GAATCCTAATAGACAAAAATGGGAAGTTGGAACTGCAGATGGCTGACAGCTTTGATGCAG
human_5+3_corrected GGATCCTAATAGACAAAAGTGGAAAATTGGAACTCCAGATGGCTGATAGTTTTGACACAG
mus_cDNA_5
rat_cDNA GTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATACAGGATAA
human_5+3_corrected GCGTATATCACTGTATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAA
mus_cDNA_5
rat_cDNA CTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGCACACAGTGGTTA
human_5+3_corrected CTGTGGTAGAACCTTTGGTCGAAGCCTATCAGGAAAATGGGATTCATCACACAGTTTTCA
mus_cDNA_5
rat_cDNA CGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAGATGCTTCTATTAGCT
human_5+3_corrected TTGGTGAAACACTTGATCTTCCATGCCATTCTACTGGTATCCCAGATGCCTCTATTAGCT
mus_cDNA_5
rat_cDNA GGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAGACAGGCAAATTCTTAACA
human_5+3_corrected GGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAAGTTCTAAACA
mus_cDNA_5

```
rat_cDNA                ATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGATCAAGGTCATTACCAATGTGTGG
human_5+3_corrected     ATGGCACATTAAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGCTGTGTGG
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                CTGCCAACCCATCAGGGGCCGACTTTTCCAGTTTTAAAGTTTCAGTTCAAAAGAAAGGCC
human_5+3_corrected     CAGCCAACCCATCAGGGGTTGATTTTTTGATTTTCCAAGTTTCAGTCAAGATGAAAGGAC
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                AAAGGATGGTTGAGCATGACAGGGAGGCAGGTGGATCTGGACTTGGAGAACCCAACTCCA
human_5+3_corrected     AAAGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTA
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                GTGTTTCCCTTAAGCAGCCAGCATCTTTGAAACTCTCTGCATCAGCTTTGACAGGGTCAG
human_5+3_corrected     TTGCTCATCTTAAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                AGGCTGGAAAACAAGTCTCCGGTGTACATAGGAAGAACAAACATAGAGACTTAATACATC
human_5+3_corrected     AGGTTGGAAAACACACCTCAAGCACAAGTAAGAGGCACAACTATCGGGAATTAACACTCC
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                GGCGGCGTGGGGATTCCACGCTCCGGCGATTCAGGGAGCATAGGAGGCAGCTCCCTCTCT
human_5+3_corrected     AGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGAATAGGAGGCATTTCCCTCCCT
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                CTGCTCGGAGAATTGACCCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATT
human_5+3_corrected     CTGCTAGGAGAATTGACCCACAACATTGGGCGGCACTGTTGGAGAAAGCTAAAAAGAATG
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                CTGTGCCAAAAAAGCAAGAAAATACCACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCG
human_5+3_corrected     CTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCACCCCCAGTGGTCACCCAACTCC
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                TGGAACTCACTGACGAGGAAAAGGATGCCTCTGGCATGATTCCTCCAGATGAAGAATTCA
human_5+3_corrected     CAAACATACCTGGTGAAGAAGACGATTCCTCAGGCATGCTCGCTCTACATGAGGAATTTA
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                TGGTTCTGAAAACTAAGGCTTCTGGTGTCCCAGGAAGGTCACCAACTGCTGACTCTGGAC
human_5+3_corrected     TGGTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAA
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                CAGTAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
human_5+3_corrected     CAATATCTGATAGTCCTATGACAAACATAAATTATGGCACAGAACTCTCCGTTGTGAATT
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                CACAAACACTACAATCTGAGCACCTTCCTGATTTCAAATTATTTAGTGTAACAAACGGTA
human_5+3_corrected     CACAAATACTACCACCTGAAGAACCCACAGATTTCAAACTGTCTACTGCTATTAAAACTA
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                CAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACAACCAACC
human_5+3_corrected     CAGCCATGTCAAAGAATATAAACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATC
mus_cDNA_5              ----------------------------------------------------------

rat_cDNA                AAAACCCAATCATTATCTTTCCATC----------AGTAGCTGAAATTCGAGATTCTG-CT
human_5+3_corrected     AACATTCATCCACTGTCTTTCCACTGCTACTTGGAGCAACTGAATTTCAGGACTCTGACA
mus_cDNA_5              ----------------------------------------------------------
```

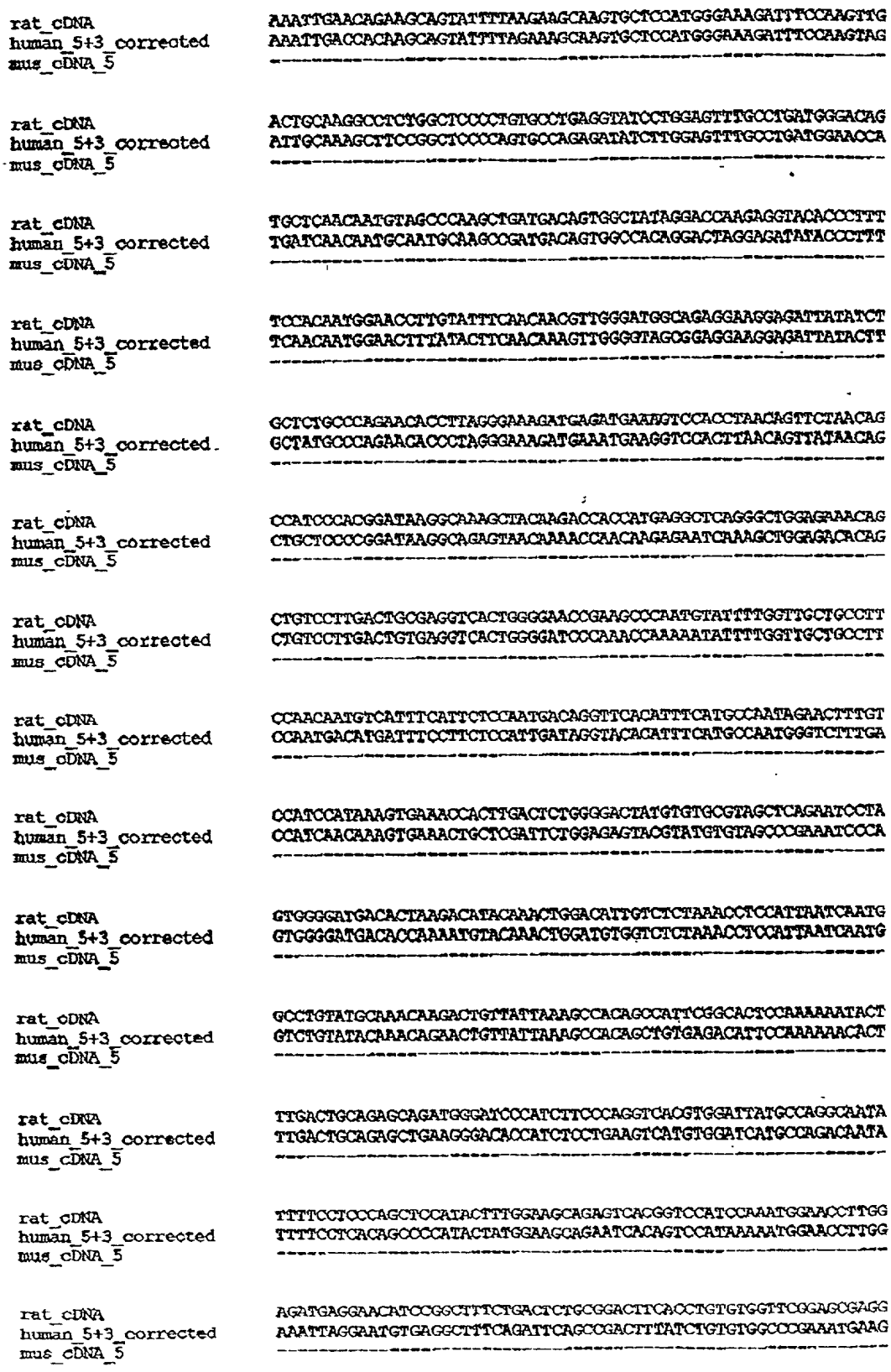

rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5
rat_cDNA
human_5+3_corrected
mus_cDNA_5

```
rat_cDNA                TTCCCAGCCCCCTCAGCACACCCTCAATACCAA------CAAGCACAAAATTCTCAAAGA
human_5+3_corrected     TTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACAAGGTTTTCAAGAA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GGAAAACTCCCTTGCACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGA
human_5+3_corrected     GGAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTAAGGA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                ATCCATATCAATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAAATAGCTCCTC
human_5+3_corrected     ATCAACATAAAGTTAGTTTACAAAAAAGCACAGCTGTGATGCTTCCTAAAACATCTCCTG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TTTTACCCACAG-GTCAGAGTTCCCCCTCAGATTCTACAACTCTCTTGACAAGTCCGCCA
human_5+3_corrected     CTTTACC-ACAGAGACAAAGTTCCCCTTTCCATTTCACCACACTTTCAACAAGTGTGATG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CCAGCTCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGT
human_5+3_corrected     CAAATTCCATCTAATACCTTGACTACCGCTCACCACACTACGACCAAA--ACACACAA-T
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GCCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAG-TGCTTCC
human_5+3_corrected     CCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCCTTC-CCACCCCTTAACCCTATGCTTCC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TAGCACCATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCCCACAGT
human_5+3_corrected     TAGTATTATAAGCAAAGACTCAAGTACAAAAAGCATCATATCAACGCAAACAGCAACCGC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GACAAGTCCTACTGCTACTGCATCTGTCATTATGTCTGAAACCCAACGAACAAGATCCAA
human_5+3_corrected     AACAACTCCTACCTTCCCTGCATCTGTCATCACTTATGAAACCCAAACAGAGAGATCTAG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                AGAAGCAAAAGACCAAATAAAGGGGCCTCG---GAAGAACAGAAACAACGCAAACACCAC
human_5+3_corrected     AGCACAAACAATACAAAGAGAAGGACCTCAAAAGAAGAACAGGACTGACCCAAACATCTC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGCTGATACCCCCTT
human_5+3_corrected     TCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACNACCTCCTNNGCTCTNNN
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GGCTTTCAGTCATTCCCCACGACAAGATGATGGTGGAAATGTAAGTGCAGTTGCTTATCA
human_5+3_corrected     NGCATTCACTCATTCCCCACCAGAAAACACAACTGGGATTTCAAGCACAATCAGTTTTCA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CTCAACAACCTCTCTTCTGGCCA---TAACTGAACTGTTTGAG-AAGTAC--ACCCAGAC
human_5+3_corrected     TTCAAGAACTCTTAATCTGACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGAC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAATCACAGGAGAGTACCAC
human_5+3_corrected     TTTGAAGAGCACAATTGCTTCTGAAACAACTTTGTCCAGCAAATCACACCAGAGTACCAC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                AGTGAAAAGAGCCTCA---GACAC---ACCACCACCACTCCTCAGCAGTGGGGCGCCCCC
human_5+3_corrected     AACTAGGAAAGCATCATTAGACACTCAACCACCACCATTCTTGAGCAGCAGTGCTACTCT
mus_cDNA_5              ------------------------------------------------------------
```

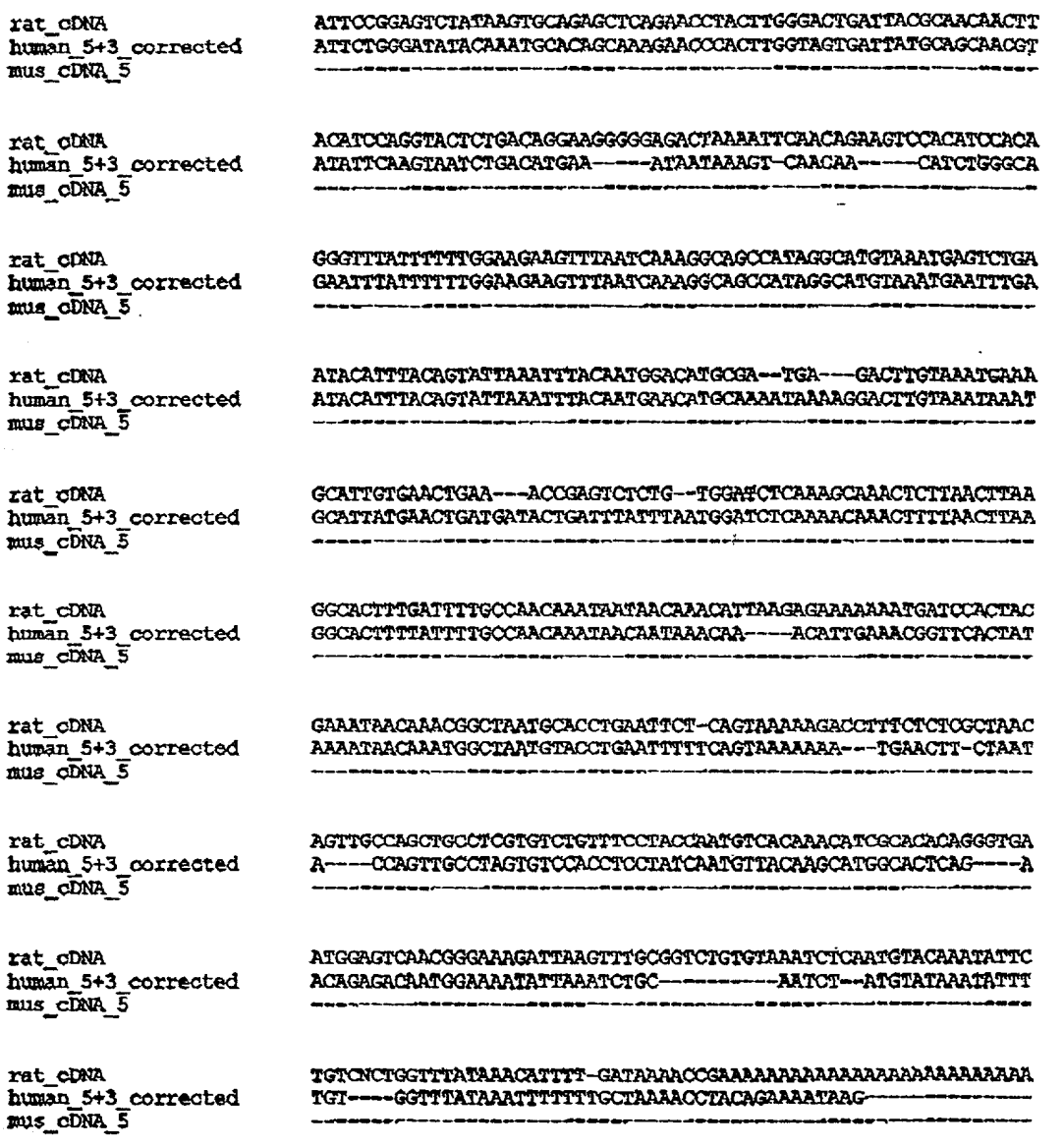

rat_cDNA
human_5+3_corrected
mus_cDNA_5
AGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTTACAGACAGCAAAGTCAC
AATGCCAGTTCCCATCTCCCCTCCCTTTACTCAGAGAGCAGTTACTGACAACGTGGCGAC
ATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCATATATGAATCTTCAAGGCACAA
TCCCATTTCCGGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGGCACAA
TACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCCCCAATCCTGAGATCATAACTGGAAC
TCCNNNNNNNCA--AATGCCAAGTTCA-------CNNAATTGNGAACCNNNNACTCNNNN
CACTGACTCTCCCTCTAATCTGTTTCCATCCACTTCTGTGCCAGCACTAAGGGTAGATAA
NACTTCATCTACNTCTAATCTGTTACATTCTACTCCCATGCCAGCACTAACAACAGTTAA
ACCACAGAATTCTAAATGGAAGCCCTCTCCCTGGCCAGAACACAAATATCAGCTCAAGTC
ATCACAGAATTCCAAATTAACTCCATCTCCCTGGGCAGAATACCAATTTTGGCAGAAACC
ATACTCCGAAACCATTGAGAAGGGCAAAAGGCCAGCAGTAAGCATGTCCCCCCACCTCAG
ATACTCAGACATTGCTGAAAAAGGCAAAAAGCCAGAAGTAAGCATGTTGGCT-ACTACAG
-CCTTCCAGAGGCCAGCACTCATGCCTCACACTGGAATACACAGAAGCATGCAGAAAAGA
GCCTGTCCGAGGCCACCACTCTTGTTTCAGATTGGGATGGACAGAAGAACACAAAGAAGA
GTGTTTTTGATAAGAAACCTGGTCAA-AACC--CAACTTCCAAACATCTGCCTTACGTCT
GTGACTTTGATAAGAAACCAGTTCAAGAAGCAACAACTTCCAAACTCCTTCCCTTTGACT
CTCTACCTAAGACTCTATTGAAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTA
CTTTGTCTAGGTATATATTTGAAAAGCCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTTA
CAGTTCCAGCTAATTCAGACGTTTTTCTTCCTTGTGAGGCTGTTGGAGACCCACTGCCCA
CTATTCCAGCTAACTCAGATGCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCCTGCCCA
TCATCCACTGGACCAGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCC
CCATTCATTGGACCAGAGTNNNNTCAGGACTTGATTTATCTAAGAGGAAACAGAATAGCA
GGTTCCACGTGCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTG
GGGTCCAGGTTCTCCCCAATGGTACCCTGTCCATCCAGAGGGTGGAAATTCAGGACCGCG
GACAGTACCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTTGT
GACAGTACTTGTGTTCCGCATCCAATCTGTTTGGCACAGACCACCTTCATGTCACCTTGT
CTGTGGTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACT
CTGTGGTTTCCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTTCATT



```
rat_cDNA                TTGGAAGTACTGTGGAACTAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTCCT
human_5+3_corrected     CCGGAAGCACTGTGGAACTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACCT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGGTCTGGG
human_5+3_corrected     GGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCTGTGG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTGGTTTTTACA
human_5+3_corrected     TGACGGTTGACGGAACATTGGTCCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                AGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCACTGTTGGTTAAGATACAAGTCATCA
human_5+3_corrected     AATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGCTGGTTAAAATACAAGTCATTG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CAGCTCCCCCTGTCATTATAGAGCAAAAGAGGCAAGCCATCGTTGGGGTTTTAGGTGGAA
human_5+3_corrected     CAGCACCACCTGTTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGCACTTGGGGTGAAA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTAGTGTTCACTGGGTCCTTT
human_5+3_corrected     GTTTAAAACTGCCCTGTACTGCAAAAGGAACTCCTCAGCCCAGCGTTTACTGGGTCCTCT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                ATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATTCCAGATTTTTCTTGTATCCAA
human_5+3_corrected     CTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCCAAGTTGTTCTTATTTTCAA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                ATGGAACTCTGTATATAAGAAGCATCGCTCCTTCAGTGAGGGGCACTTATGAGTGCATTG
human_5+3_corrected     ATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACTTATGAATGCATTG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAGTGATTCTTACTGTGGAAGAGGGAGAGA
human_5+3_corrected     CTACCAGTTCCACTGGTTCGGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CAATCCCCAGGATAGAAACTGCCTCTCAGAAATGGACTGAGGTGAATTTGGGTGAGAAAT
human_5+3_corrected     CCAGCCCCAGGATAGAAGCTGCATCCCAGAAAAGGACTGAAGTGAATTTTGGGGAGAAAT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TACTACTGAACTGCTCAGCTACTGGGGATCCAAAGCCTAGAATAATCTGGAGGCTGCCAT
human_5+3_corrected     TACTACTGAACTGCTCAGCCACTGGGGAGCCCAAACCCCAAATAATGTGGAGGTTACCAT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CCAAGGCTGTCATCGACCAGTGGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATG
human_5+3_corrected     CCAAGGCTGTGGTCGACCAGTGG----------GCAGCTGGATCCACGTCTACCCTAATG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GATCCTTGGTGGTTGGGTCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAA
human_5+3_corrected     GATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTCTACTTGTGTGTGGCAA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GAAACAAAATGGGAGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCA
human_5+3_corrected     GAAACAAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTGAAACCTGCCA
mus_cDNA_5              ------------------------------------------------------------
```

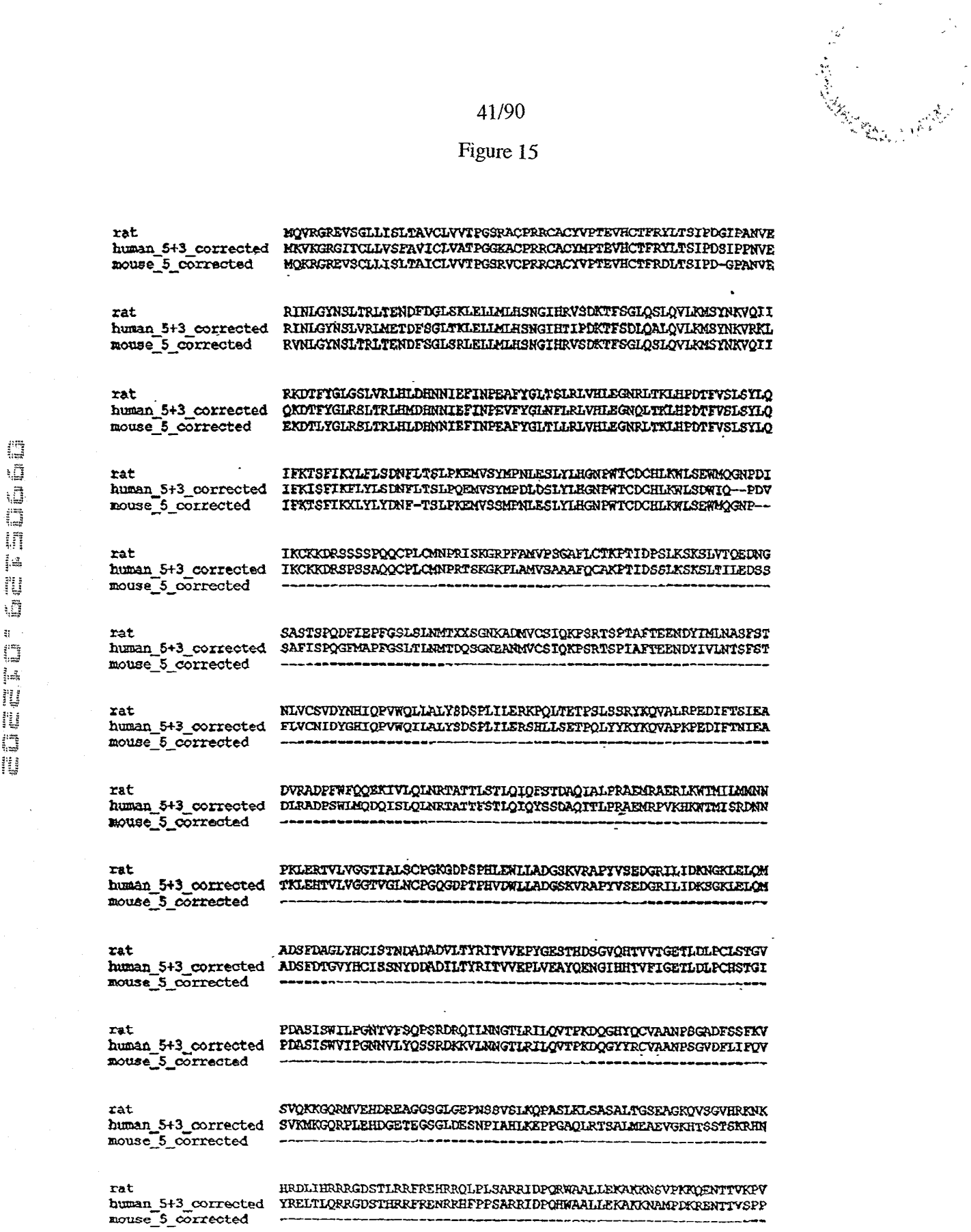

rat_cDNA
human_5+3_corrected
mus_cDNA_5
AAATTGAACAGAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTG
AAATTGACCACAAGCAGTATTTTAGAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTAG
ACTGCAAGGCCTCTGGCTCCCCTGTGCCTGAGGTATCCTGGAGTTTGCCTGATGGGACAG
ATTGCAAAGCTTCCGGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCA
TGCTCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCTTT
TGATCAACAATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCCTTT
TCCACAATGGAACCTTGTATTTCAACAACGTTGGGATGGCAGAGGAAGGAGATTATATCT
TCAACAATGGAACTTTATACTTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTT
GCTCTGCCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGTTCTAACAG
GCTATGCCCAGAACACCCTAGGGAAAGATGAAATGAAGGTCCACTTAACAGTTATAACAG
CCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGGCTGGAGAAACAG
CTGCTCCCCGGATAAGGCAGAGTAACAAAACCAACAAGAGAATCAAAGCTGGAGACACAG
CTGTCCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTATTTTGGTTGCTGCCTT
CTGTCCTTGACTGTGAGGTCACTGGGGATCCCAAACCAAAAATATTTTGGTTGCTGCCTT
CCAACAATGTCATTTCATTCTCCAATGACAGGTTCACATTTCATGCCAATAGAACTTTGT
CCAATGACATGATTTCCTTCTCCATTGATAGGTACACATTTCATGCCAATGGGTCTTTGA
CCATCCATAAAGTGAAAACCACTTGACTCTGGGGACTATGTGTGCGTAGCTCAGAATCCTA
CCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCA
GTGGGGATGACACTAAGACATACAAACTGGACATTGTCTCTAAACCTCCATTAATCAATG
GTGGGGATGACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAATG
GCCTGTATGCAAACAAGACTGTTATTAAAGCCACAGCCATTCGGCACTCCAAAAAATACT
GTCTGTATACAAACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAACACT
TTGACTGCAGAGCAGATGGGATCCCATCTTCCCAGGTCACGTGGATTATGCCAGGCAATA
TTGACTGCAGAGCTGAAGGGACACCATCTCCTGAAGTCATGTGGATCATGCCAGACAATA
TTTTCCTCCCAGCTCCATACTTTGGAAGCAGAGTCACGGTCCATCCAAATGGAACCTTGG
TTTTCCTCACAGCCCCATACTATGGAAGCAGAATCACAGTCCATAAAAATGGAACCTTGG
AGATGAGGAACATCCGGCTTTCTGACTCTGCGGACTTCACCTGTGTGGTTCGGAGCGAGG
AAATTAGGAATGTGAGGCTTTCAGATTCAGCCGACTTTATCTGTGTGGCCCGAAATGAAG

```
rat_cDNA                GAGGAGAGAGTGTGTTGGTAGTGCAGTTAGAAGTCCTAGAAATGCTGAGAAGACCAACAT
human_5+3_corrected     GTGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGACCGACAT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TCAGAAACCCATTCAACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACT
human_5+3_corrected     TTAGAAATCCATTTAATGAAAAAATAGTTGCCCAGCTGGGAAAGTCCACAGCATTGAATT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GCTCTGTGGATGGGAACCCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGT
human_5+3_corrected     GCTCTGTTGATGGTAACCCACCACCTGAAATAATCTGGATTTTACCAAATGGCACACGAT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TTGCTAACAGACCACACAATTCCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTT
human_5+3_corrected     TTTCCAATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATTT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                ACAAAGCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT
human_5+3_corrected     CTAAAACAACTCGGGAGGATGCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                ACATCGAGAAACTCATCCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGAAC
human_5+3_corrected     ATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCAC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTGTCTGATGGGA
human_5+3_corrected     CAGGGACAGTAAAAGGCATCAGTGGAGAATCTCTATCACTGCATTGTGTGTCTGATGGAA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGACAGGCCTCAAG
human_5+3_corrected     TCCCTAAGCCAAATATCAAATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCCTCAAA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TGGATGGAAAATACATACTGCATGAAAATGGCACGCTGGTCATCAAAGCAACAACAGCTC
human_5+3_corrected     TTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCTT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                ACGACCAAGGAAATTATATCTGTAGGGCTCAAAACAGTGTTGGCCAGGCAGTTATTAGCG
human_5+3_corrected     ATGACAGAGGAAACTATATCTGTAAGGCTCAAAATAGTGTTGGTCATACACTGATTACTG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TGTCAGTGATGGTTGTGGCCTACCCTCCCCGAATCATAAACTACCTACCCAGGAACATGC
human_5+3_corrected     TTCCAGTAATGATTGTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGTGTGGCCTTGGGAATCCCCAAGCCAA
human_5+3_corrected     TCACCAGGACAGGGGCAGCCTTTCAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                AAGTCACCTGGGAGACGCCAAGACACTCCCTGCTCTCAAAAGCAACAGCAAGAAAACCCC
human_5+3_corrected     AAATCACATGGGAGATGCCTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACAC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                ATAGAAGTGAGATGCTTCACCCACAAGGTACGCTGGTCATTCAGAATCTCCAAACCTCGG
human_5+3_corrected     ATGGAAGTGAGCAGCTTCACTTACAAGGTACCCTAGTCATTCAGAATCCCCAAACCTCCG
mus_cDNA_5              ------------------------------------------------------------
```



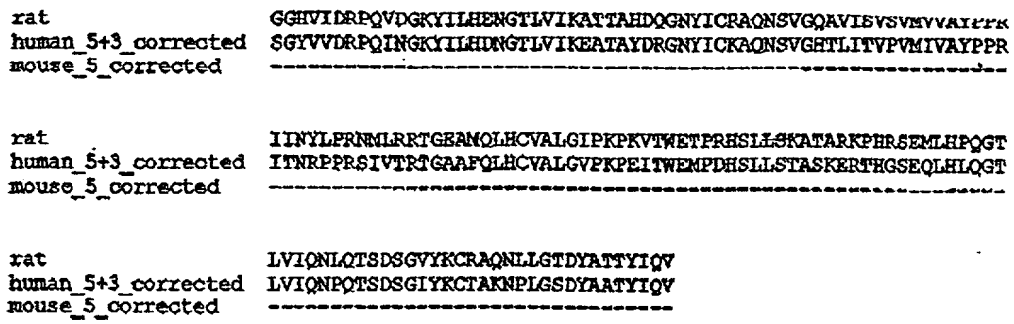
(rat_cDNA: SEQ ID NO:7)
(human_5+3 corrected: SEQ ID NO:8)
(mus_cDNA_5: SEQ ID NO:9)

Figure 15


rat MQVRGREVSGLLISLTAVCLVVTPGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
human_5+3_corrected MKVKGRGITCLLVSFAVICLVATPGGKACPRRCACYMPTEVHCTFRYLTSIPDSIPPNVE
mouse_5_corrected MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYVPTEVHCTFRDLTSIPD-GPANVE
rat RINLGYNSLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
human_5+3_corrected RINLGYNSLVRLMETDFSGLTKLELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRKL
mouse_5_corrected RVNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
rat RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQ
human_5+3_corrected QKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYLQ
mouse_5_corrected EKDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTFVSLSYLQ
rat IFKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESLYLHGNPWTCDCHLKWLSEWMQGNPDI
human_5+3_corrected IFKISFIKFLYLSDNFLTSLPQEMVSYMPDLDSLYLHGNPWTCDCHLKWLSDWIQ--PDV
mouse_5_corrected IFKTSFIKXLYLYDNF-TSLPKEMVSSMPNLESLYLHGNPWTCDCHLKWLSEWMQGNP--
rat IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNG
human_5+3_corrected IKCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSS
mouse_5_corrected ----------
rat SASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFST
human_5+3_corrected SAFISPQGFMAPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPIAFTEENDYIVLNTSFST
mouse_5_corrected ----------
rat NLVCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEA
human_5+3_corrected FLVCNIDYGHIQPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEA
mouse_5_corrected ----------
rat DVRADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNN
human_5+3_corrected DLRADPSWLMQDQISLQLNRTATTFSTLQIQYSSDAQITLPRAEMRPVKHKWTMISRDNN
mouse_5_corrected ----------
rat PKLERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQM
human_5+3_corrected TKLEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
mouse_5_corrected ----------
rat ADSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
human_5+3_corrected ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIHHTVFIGETLDLPCHSTGI
mouse_5_corrected ----------
rat PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKDQGHYQCVAANPSGADFSSFKV
human_5+3_corrected PDASISWVIPGNNVLYQSSRDKKVLNNGTLRILQVTPKDQGYYRCVAANPSGVDFLIFQV
mouse_5_corrected ----------
rat SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKLSASALTGSEAGKQVSGVHRKNK
human_5+3_corrected SVKMKGQRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTSALMEAEVGKHTSSTSKRHN
mouse_5_corrected ----------
rat HRDLIHRRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTVKPV
human_5+3_corrected YRELTLQRRGDSTHRRFRENRRHFPPSARRIDPQHWAALLEKAKKNAMPDKRENTTVSPP
mouse_5_corrected ----------

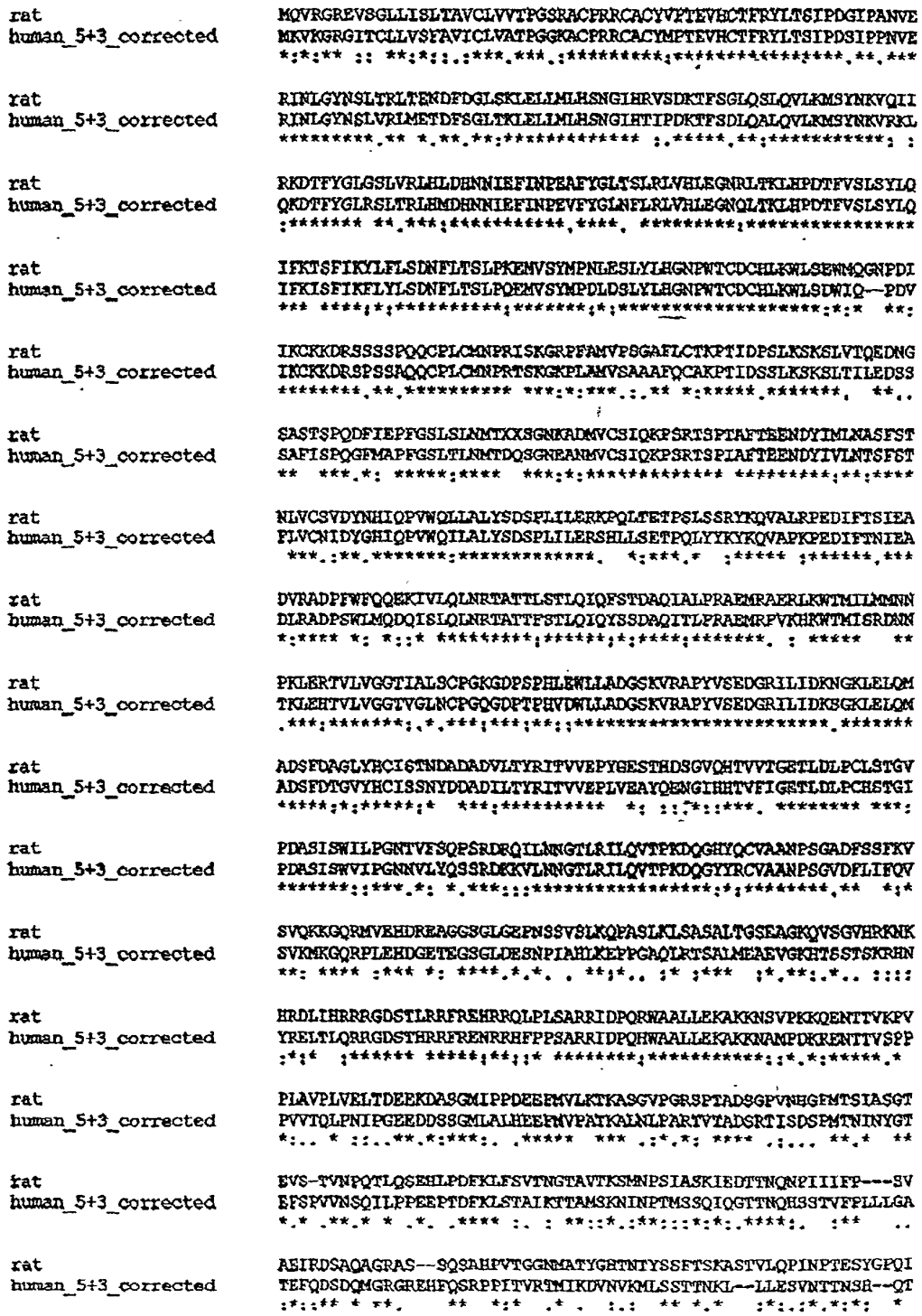
```
rat                   PLAVPLVELTDEEKDASGMIPPDEEPMVLKTKASGVPGRSPTADSGPVNHGPMTSIASGT
human_5+3_corrected   PVVTQLPNIPGEEDDSSGMLALHEEPMVPATKALNLPARTVTADSRTISDSPMTNINYGT
mouse_5_corrected     ------------------------------------------------------------

rat                   EVS-TVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIIPP----SV
human_5+3_corrected   EFSPVVNSQILPPEEPTDFKLSTAIKTTAMSKNINPTMSSQIQGTTNQHSSTVFPLLLGA
mouse_5_corrected     ------------------------------------------------------------

rat                   AEIRDSAQAGRAS--SQSAHPVTGGNMATYGHTNTYSSFTSKASTVLQPINPTKSYGPQI
human_5+3_corrected   TEFQDSDQMGRGREHFQSRPPITVRTMIKDVNVKMLSSTTNKL--LLESVNTTNSH--QT
mouse_5_corrected     ------------------------------------------------------------

rat                   PITGVSRPSSSDISSHTTADPSFSSHPSGSHTTASSLPHIPRNNNTGNFPLSRHLGRERT
human_5+3_corrected   SVREVSEPRHNHFYSHTTQILSTSTFPSDPHTAAHSQFPIPRNS-TVNIPLFRRFGRQRK
mouse_5_corrected     ------------------------------------------------------------

rat                   IWSRGRVENPHRTPVLRRHRHRTVRPAIKGPANKNVSQVPATEYPGMCHTCPSAEGLTVA
human_5+3_corrected   IGGRGRIISPYRTPVLRRHRYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRERLTTA
mouse_5_corrected     ------------------------------------------------------------

rat                   TAALSVPSSSHSALPKTNNVGVIAEESTTVVKKPLLLFKDKQNVDIEIITTTTTKYSGGES
human_5+3_corrected   TAALSFPSAAPITFPKADIARVPSEESTTLVQNPLLLLENKP--SVEKTTPTIKYFRTEI
mouse_5_corrected     ------------------------------------------------------------

rat                   NHVIPTEASMTSAPTSVSLGKSPVDNSGHLSMPGTIQTGKDSVETTPLPSPLSTP--SIP
human_5+3_corrected   SQVTPTGAVMTYAPTSIPMEKTHKVNASYPRVSSTNEAKRDSVITSSLSGAITKPPMTII
mouse_5_corrected     ------------------------------------------------------------

rat                   TSTKFSKRKTPLHQIFVNWQKEGMLKNPYQFGLQKNPAAKLPKIAPLLPTGQSSPSDST
human_5+3_corrected   AITRFSRRKIPWQQNFVNNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFT
mouse_5_corrected     ------------------------------------------------------------

rat                   TLLTSPPPALSTTMAATQNKGTEVVSGARSLSAGKKQ-PFTNSSPVLPSTISKRSNTLNF
human_5+3_corrected   TISTSVMQIPSNTLTTAHHTTTKTHNPG-SLPT-KKELPFPPLNPMLPSIISKDSSTKSI
mouse_5_corrected     ------------------------------------------------------------

rat                   LSTETPT-VTSPTATASVIMSETQRTRSKEAKDQIKG-P-RKNRNNANTTPRQVSGYSAY
human_5+3_corrected   ISTQTAIPATTPTFPASVITYETQTERSRAQTIQREQEPQKKNRTDPNISPDQSSGFTTP
mouse_5_corrected     ------------------------------------------------------------

rat                   SALTTADTPLAFSHSPRQDDGGNVSAVAYHSTTS--LLAITELFEKYTQTLGNTTALETT
human_5+3_corrected   TAMTP---PALAFTHSPPENTTGISSTISFHSRTLNLTDVIEELAQASTQTLKSTIASETT
mouse_5_corrected     ------------------------------------------------------------

rat                   LLSKSQESTTVKRAS-DTP-PFLLSSGAPPVPTPSPPPFTKGVVTDSKVTSAFQMTSNRV
human_5+3_corrected   LSSKSHQSTTTRKASLDTPIPPFLSSSATIMPVPISPPFTQRAVTDTRGDSHFRLMTNTV
mouse_5_corrected     ------------------------------------------------------------

rat                   VTIYESSRHNTDLQQPSAEASPNPEIITGTTDSPSNLFPSTSVPALRVDKPQNSKWKPSP
human_5+3_corrected   VKLHESSRHN--LQMPSSQLEP--------LTSSTSNLLHSTPMPALTTVKSQNSKLTPSP
mouse_5_corrected     ------------------------------------------------------------

rat                   WPEHKYQLKSYSETIEKGKRPAVSMSPHLSLPEASTHASHWNTQKHAEKSVFDKKPGQNP
human_5+3_corrected   WAEYQFWHKPYSDIAEKGKKPEVSMLATTGLSEATTLVSDWDGQKNTKKSDFDKKPVQEA
mouse_5_corrected     ------------------------------------------------------------
```

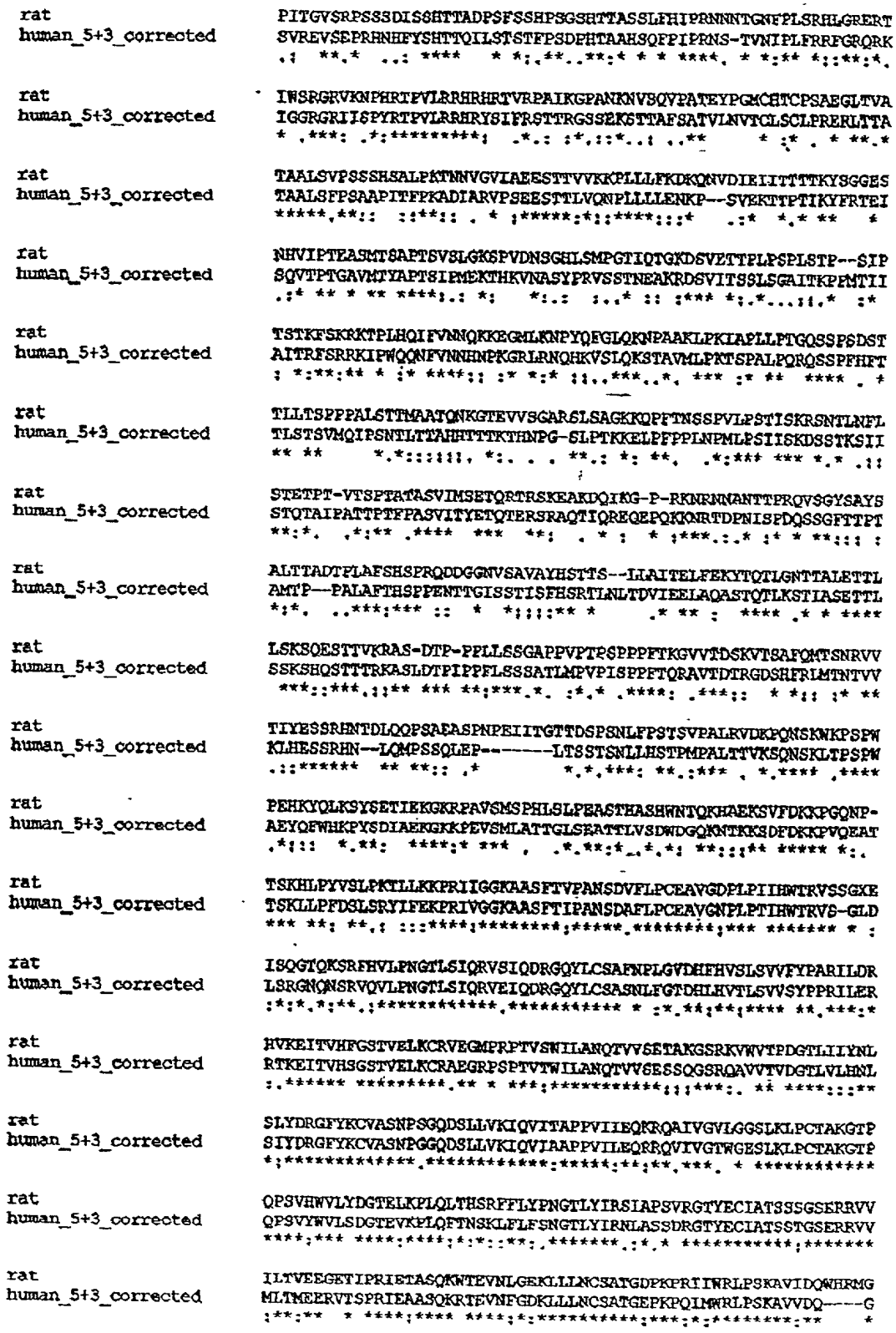
```
rat                     -TSKHLPYVSLPKTLLKKPRIIGGKAASFTVPANSDVFLPCEAVGDPLPIIHWTRVSSGX
human_5+3_corrected     TTSKLLPFDSLSRYIFEKPRIVGGKAASFTIPANSDAFLPCEAVGNPLPTIHWTRVS-GL
mouse_5_corrected       ------------------------------------------------------------

rat                     EISQGTQKSRFHVLPNGTLSIQRVSIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILD
human_5+3_corrected     DLSRGNQNSRVQVLPNGTLSIQRVEIQDRGQYLCSASNLFGTDHLHVTLSVVSYPPRILE
mouse_5_corrected       ------------------------------------------------------------

rat                     RHVKEITVHPGSTVELKCRVEGMPRPTVSWILANQTVVSETAKGSRKVWVTPDGTLIIYN
human_5+3_corrected     RRTKEITVHSGSTVELKCRAEGRPSPTVTWILANQTVVSESSQGSRQAVVTVDGTLVLHN
mouse_5_corrected       ------------------------------------------------------------

rat                     LSLYDRGFYKCVASNPSGQDSLLVKIQVIYAPPVIIEQKRQAIVGVLGGSLKLPCTAKGT
human_5+3_corrected     LSIYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVILEQRRQVIVGTWGESLKLPCTAKGT
mouse_5_corrected       ------------------------------------------------------------

rat                     PQPSVHWVLYDGTELKPLQLTHSRFFLYPMGTLYIRSIAPSVRGIYSCIATSSSGSERRV
human_5+3_corrected     PQPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLASSDRGTYECIATSSTGSERRV
mouse_5_corrected       ------------------------------------------------------------

rat                     VILTVEEGETIPRIETASQKWTEVNLGEKLLLNCSATGDPKPRIIWRLPSKAVIDQWHRM
human_5+3_corrected     VMLTMEERVTSPRIEAASQKRTEVNFGDKLLLNCSATGEPKPQIMWRLPSKAVVDQ-----
mouse_5_corrected       ------------------------------------------------------------

rat                     GSRIHVYPNGSLVVGSVTEKDAGDYLCVARNKMGDDLVLMHVRLRLTPAKIEQKQYFKKQ
human_5+3_corrected     GSWIHVYPNGSLFIGSVTEKDSGVYLCVARNKMGDDLILMHVSLRLKPAKIDHKQYFRKQ
mouse_5_corrected       ------------------------------------------------------------

rat                     VLHGKDFQVDCKASGSPVPEVSWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFNNV
human_5+3_corrected     VLHGKDFQVDCKASGSPVPEISWSLPDGTNINNAMQADDSGHRTRRYTLFNNGTLYFNKV
mouse_5_corrected       ------------------------------------------------------------

rat                     GMAEEGDYICSAQNTLGKDEMKVHLTVLTAIPRIRQSYKTTMRLRAGETAVLDCEVTGEP
human_5+3_corrected     GVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIRQSNKTNKRIKAGDTAVLDCEVTGDP
mouse_5_corrected       ------------------------------------------------------------

rat                     KPNVFWLLPSNNVISFSNDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSGDDTKTYKLD
human_5+3_corrected     KPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSGEYVCVARNPSGDDTKMYKLD
mouse_5_corrected       ------------------------------------------------------------

rat                     IVSKPPLINGLYANKTVIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSR
human_5+3_corrected     VVSKPPLINGLYTNRTVIKATAVRHSKKHFDCRAEGTPSPEVMWIMPDNIFLTAPYYGSR
mouse_5_corrected       ------------------------------------------------------------

rat                     VTVHPNGTLEMRNIRLSDSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIA
human_5+3_corrected     ITVHKNGTLEIRNVRLSDSADFICVARNEGGESVLVVQLEVLEMLRRPTFRNPFNEKIVA
mouse_5_corrected       ------------------------------------------------------------

rat                     QAGKPVALNCSVDGNPPPEITWILPDGTQFANRPHNSPYIMAGNGSLILYKATRNKSGKY
human_5+3_corrected     QLGKSTALNCSVDGNPPPEIIWILPNGTRFSNGPQSYQYLIASNGSFIISKTTREDAGKY
mouse_5_corrected       ------------------------------------------------------------

rat                     RCAARNKVGYIEKLILLEIGQKPVILTYEPGMVKSVSGEPLSLHCVSDGIPKPNVKWTTP
human_5+3_corrected     RCAARNKVGYIEKLVILEIGQKPVILTYAPGTVKGISGESLSLHCVSDGIPKPNIKWTMP
mouse_5_corrected       ------------------------------------------------------------
```

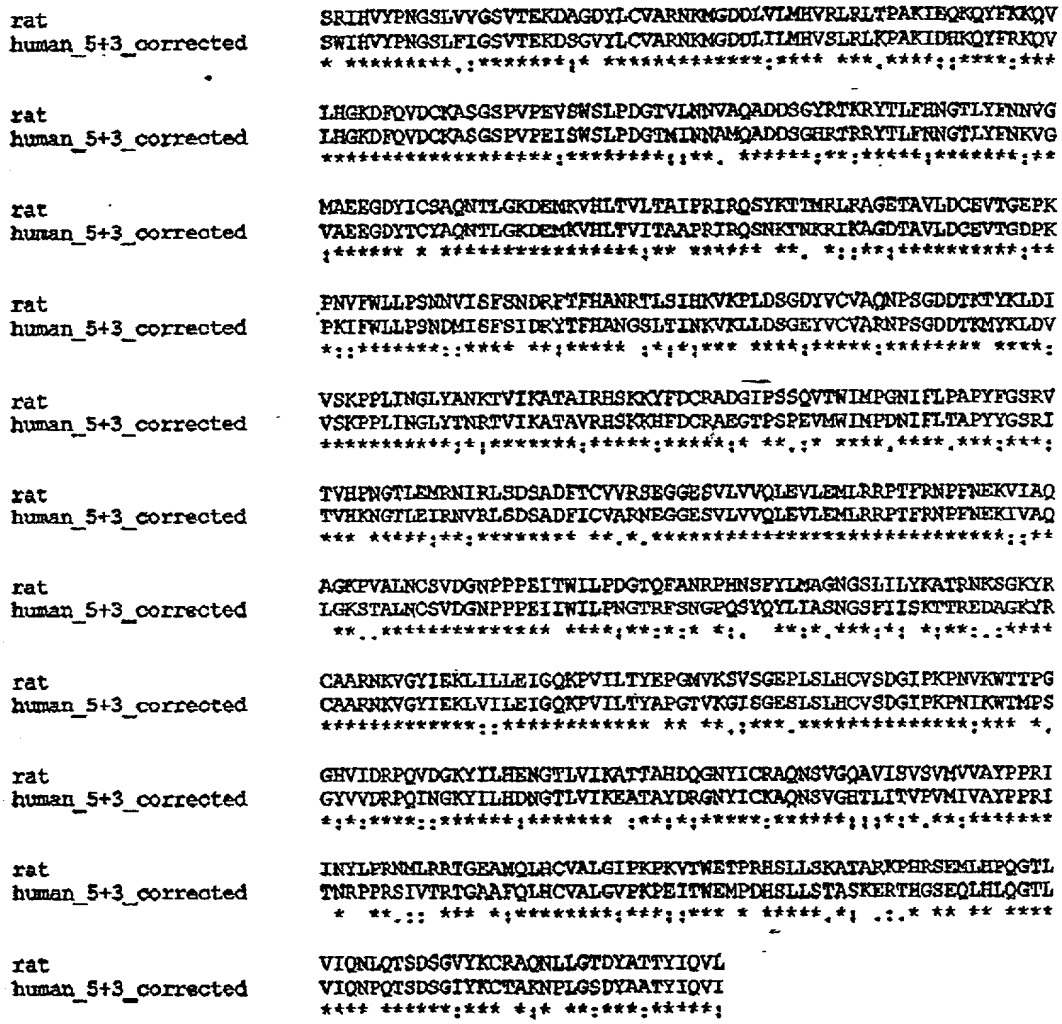


(rat: SEQ ID NO:10)
(human_5+3_corrected: SEQ ID NO:11)
(mouse_5_corrected: SEQ ID NO:12)

Figure 16




rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected
rat
human_5+3_corrected



(rat: SEQ ID NO:13)
(human_5+3_corrected: SEQ ID NO:14)

Figure 17

MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYVPTEVHCTFRDLTSIPDGPANVER
VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQIIE
KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTFVSLSYLQIF
KTSFIKXLYLYDNFTSLPKEMVSSMPNLESLYLHGNPWTCDCHLKWLSEWMQGNP
(SEQ ID NO: 15)

Figure 18

MKVKGRGITCLLVSFAVICLVATPGGKACPRRCACYMPTEVHCTFRYLTSIPDSIPPNVE
RINLGYNSLVRLMETDFSGLTKLELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK
LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYL
QIFKISFIKFLYLSDNFLTSLPQEMSYMPDLDSLLYLHGNPWTCDCHLRWLSDWIQPDVI
KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS
AFISPQGFMAPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPIAFTEENDYIVLNTSFSTFL
VCNIDYGHIQPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEADLR
ADPSWLMQDQISLQLNRTATTFSTLQIQYSSDAQITLPRAEMRPVKHKWTMISRDNNTK
LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIHHTVFIGETLDLPCHSTGIP
DASISWVIPGNNVLYQSSRDKKVLNNGTLRILQVTPKDQGYYRCVAANPSGVDFLIFQV
SVKMKGQRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTSALMEAEVGKHTSSTSKRH
NYRELTLQRRGDSTHRRFRENRRHFPPSARRIDPQHWAALLEKAKKNAMPDKRENTTV
SPPPVVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTADSRTISDSPMTNIN
YGTEFSPVVNSQILPPEEPTDFKLSTAILTTAMSKNINPTMSSQIQGTTNQHSSTVFPLLLG
ATEFQDSDQMGRGREHFQSRPPITVRTMIKDVNVKMLSSTTNKLLLESVNSHQTSVREV
SEPRHNHFYSHTTQILSTSTFPSDPHTAAHSQFPIPRNSTVNIPLFRRFGRQRKIGGRGRIIS
PYRTPVLRRHRYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRELTTATAALSFPSAAPI
TFPKADIARVPSEESTTLVQNPLLLLENKPSVEKTTPTIKYFRTEISQVTPTGAVMTYAPT
SIPMEKTHKVNASYPRVSSTNEAKRDSVITSSLSGAITKPPMTIIAITRFSRRKIPWQQNFV
NNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFRRLSTSVMQIPSNTLTT
AHHTTTKTHNPGSLPTKKELPFPPLNPMLPSIISKDSSTKSIISTQTAIPATTPTFPASVITYE
TQTERSRAQTIQREQEPQKKNRTDPNISPDQSSGFTTPTAMTPPALAFTHSPPENTTGISST
ISFHSRTLNLTDVIEELAQASTQTLKSTIASETTLSSKSHQSTTTRKASLDTPIPPFLSSSAT
LMPVPISPPFTQRAVTDTRGDSHFRLMTNTVVKLHESSRHNLQMPSSQLEPLTSSTSNLL
HSTPMPALTTVKSQNSKLTPSPWAEQFWHKPYSDIAEKGKKPEVSMLATTGLSEATTLV
SDWDGQKNTKKSDFDKKPVQEATTSKLLPFDSLSRYIFEKPRIVGGKAASFTIPANSDAF
LPCEAVGNPLPTIHWTRVSGLDLSRGNQNSRVQVLPNGTLSIQRVEIQDRFQYLCSASNL
FGTDHLHVTLSVVSYPPRILERRTKEITVHSGSTVELKCRAEGRPSPTVTWILANQTVVSE
SSQGSRQAVVTVDGTLVLHNLSIYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVILEQRR
QVIVGTWGESLKLPCTAKGTPQPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLAS

SDRFTYECIATSSTGSERRVVMLTMEERVTSPRIEAASQKRTEVNFGDKLLLNCSATGEP
KPQIMRLPSKAVVDQGSWIHYPNGSLFIGVTEKDSGVYLCVARNKMGDDLILMHVSLR
LKPAKIDHKQYFRKQVLHGKDFQVDCKASGSPVPEISWSLPDGTMINNAMQADDSGHR
TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIRQSNKTN
KRIKAGDTAAVLDCEVTGDPKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG
EYVCVARNPSGDDTKMYKLDVVSKPPLINGLYTNRTVIKATAVRHSKKHFDCRAEGTP
SPEVMWIMPDNIFLTAPYYGSRITVHKNGTLEIRNVRLSADFICVARNEGGESVLVVQLE
VLEMLRRPTFRNPRNPFNEKIVAQLGKSTALNCVSVDGNPPPEIIWILPNGTRFSNGPQSY
QYLIASNGSFIISKTTREDAGKYRCAARNKVGYIEKLVILEIGQKPVILTYAPGTVKGISGE
SLSLHCVSDGIPKPNIKWTMPSGYVVDRPQINGKYILHDNGTLVIKEATAYDRGNYICKA
QNSVGHTLITVPVMIVAYPPRITNRPPRSIVTRTGAAFQLHCVALGVPKPEITWEMPDHS
LLSTASKERTHGSEQLHLQGTLVIQNPQTSDSGIYKCTAKNPLGSDYAATYIQVI (SEQ ID
NO: 16)

Figure 19

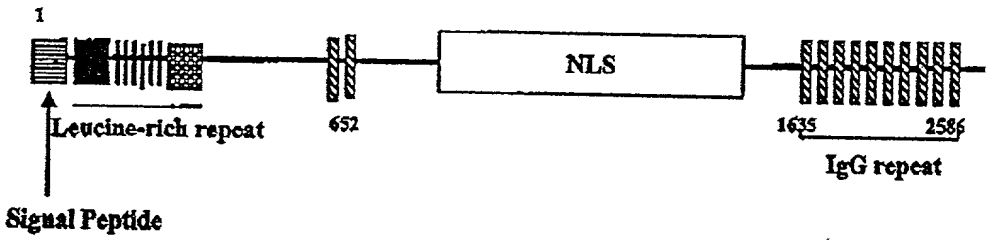
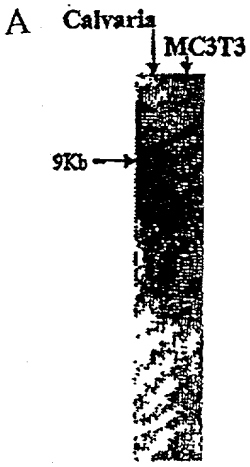


Figure 20

A



B

Cell lines & primary cells		
C6	Rat glioma	-
C2C12	Myoblasts	-
STO	Fibroblasts	-
Bone marrow	Stromal cells	- (+)
PC-3	prostatic carcinoma	-
SAOS-2	osteosarcoma	-
U2OS	osteosarcoma	+
Articular cartilage	chondrocytes	-
ROS	osteosarcoma	-
Calvaria	Osteoblasts	+
C3H10T1/2	mesenchyme	- (+)
MC3T3-E1	preosteoblastic	- (+)

Figure 21

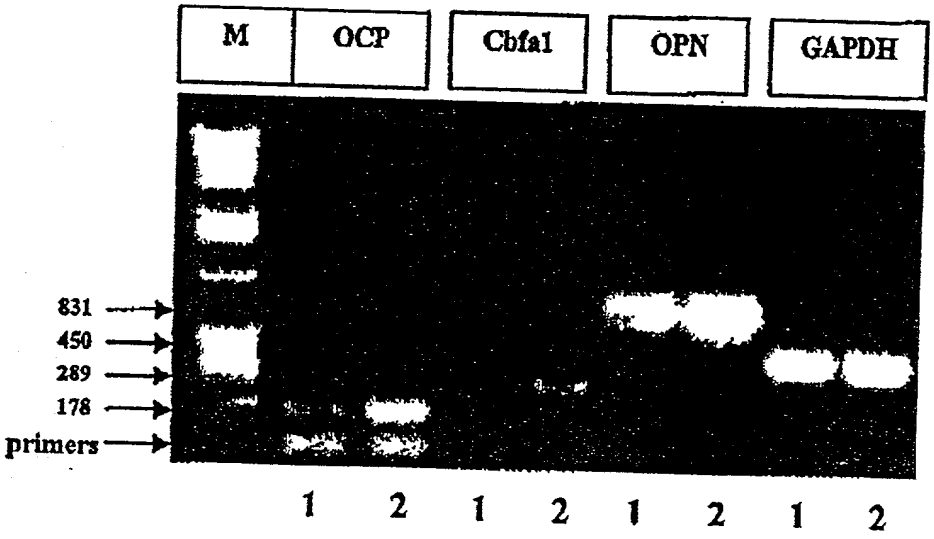


Figure 22

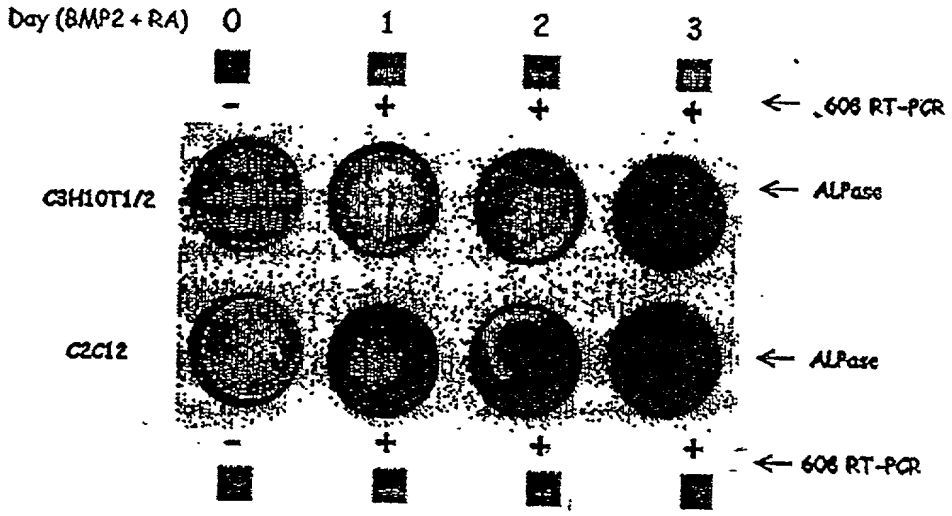


Figure 23

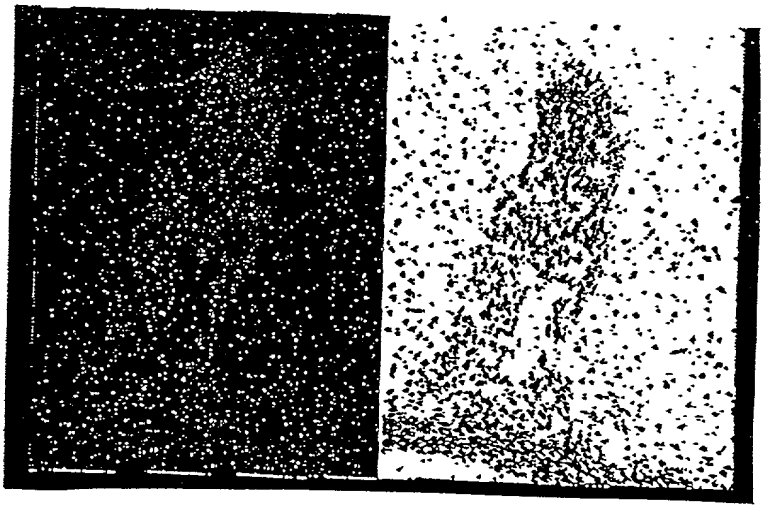


Figure 24

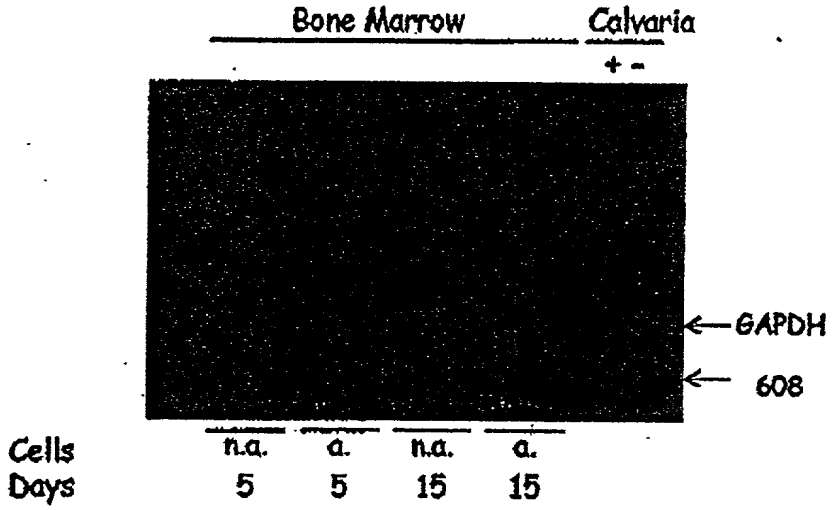


Figure 25

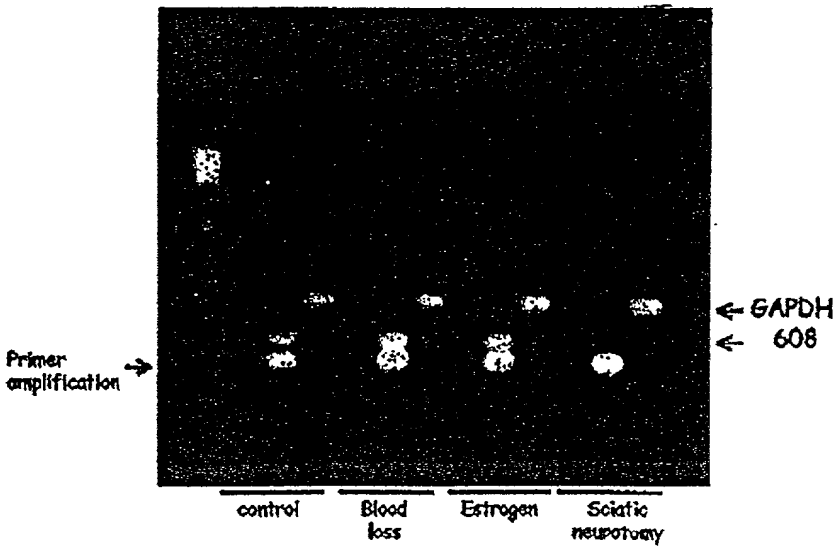


Figure 26

Figure 27

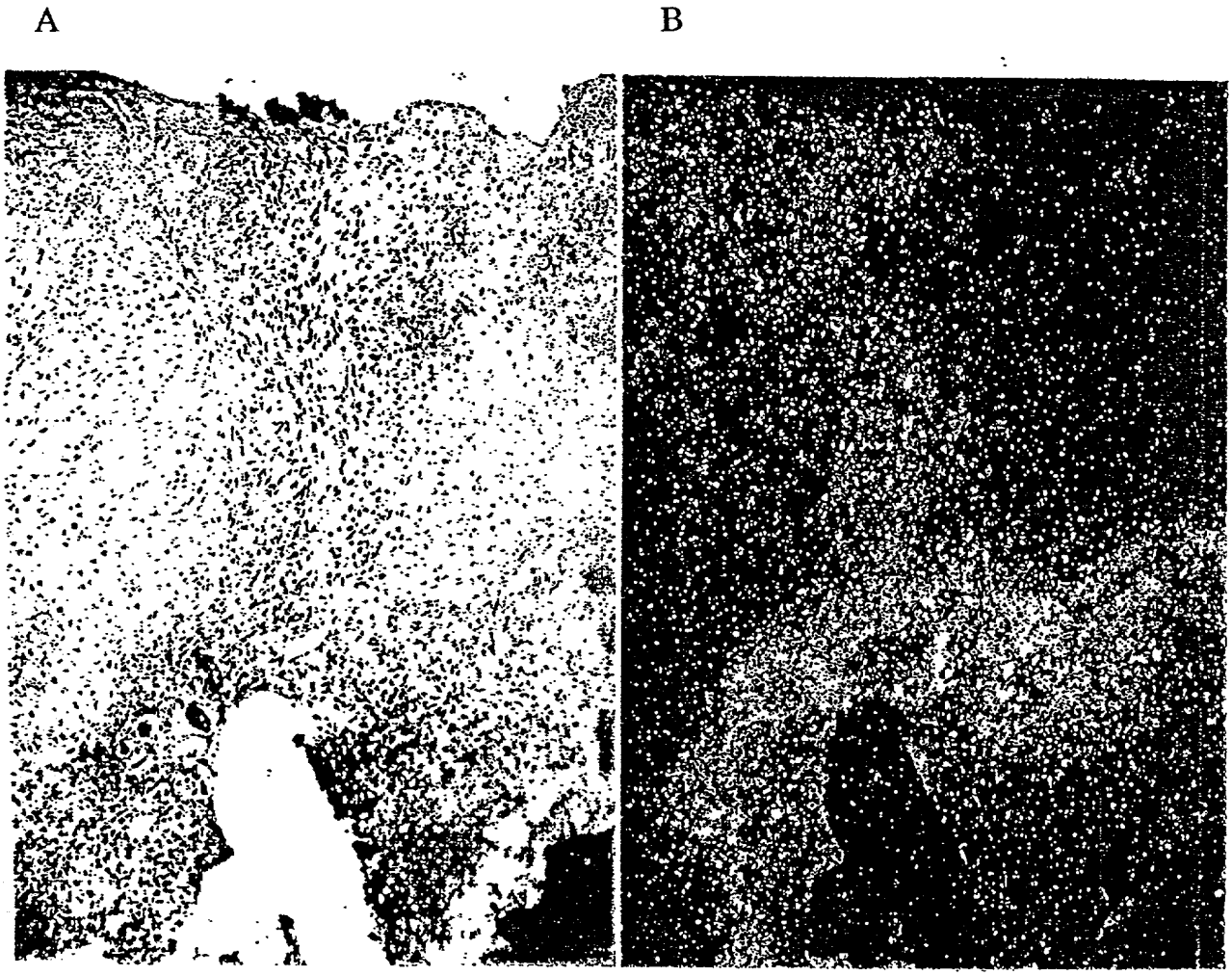
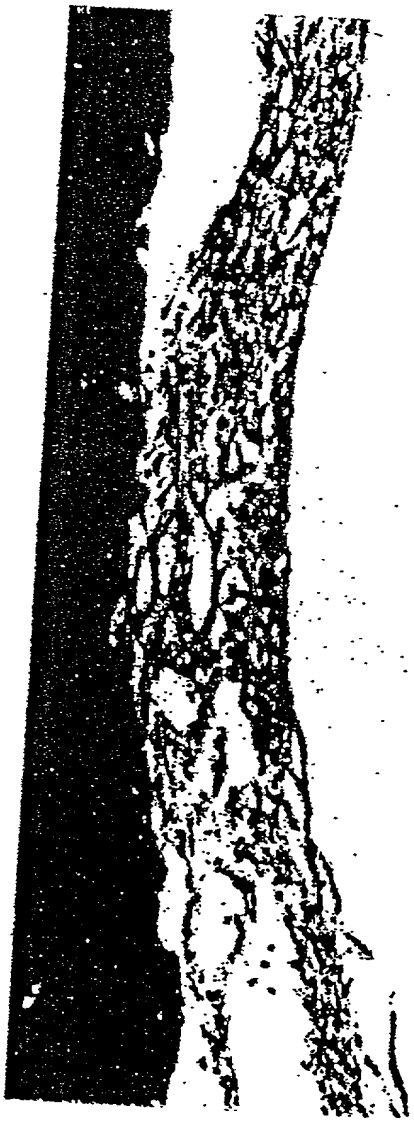


Figure 28

Figure 29

A



B

Figure 30

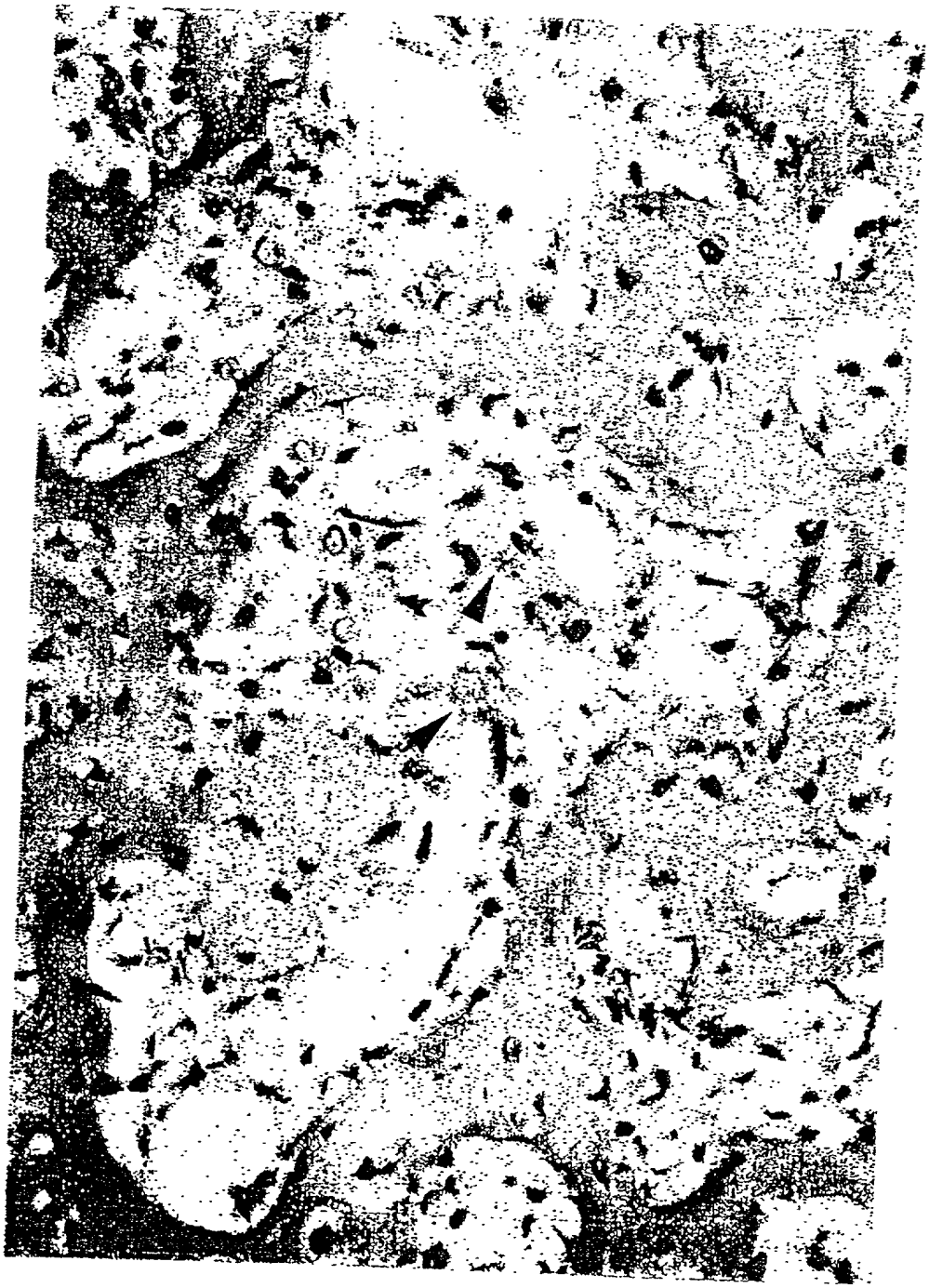
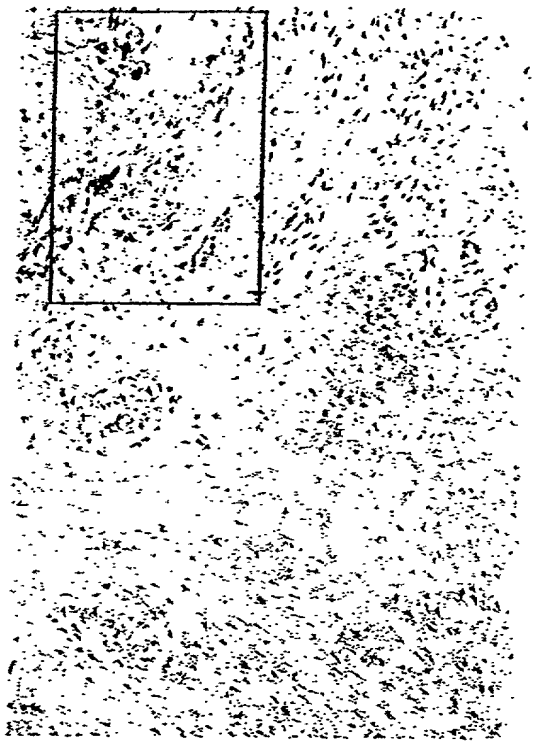


Figure 31

Figure 32

A



B

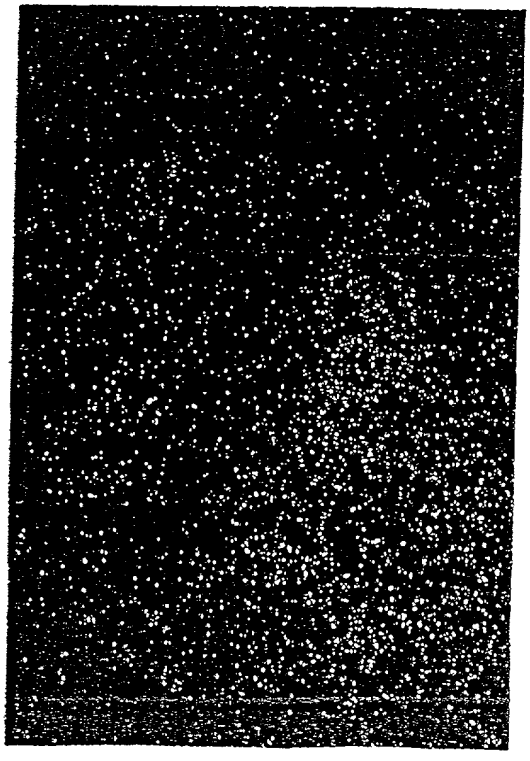


Figure 33

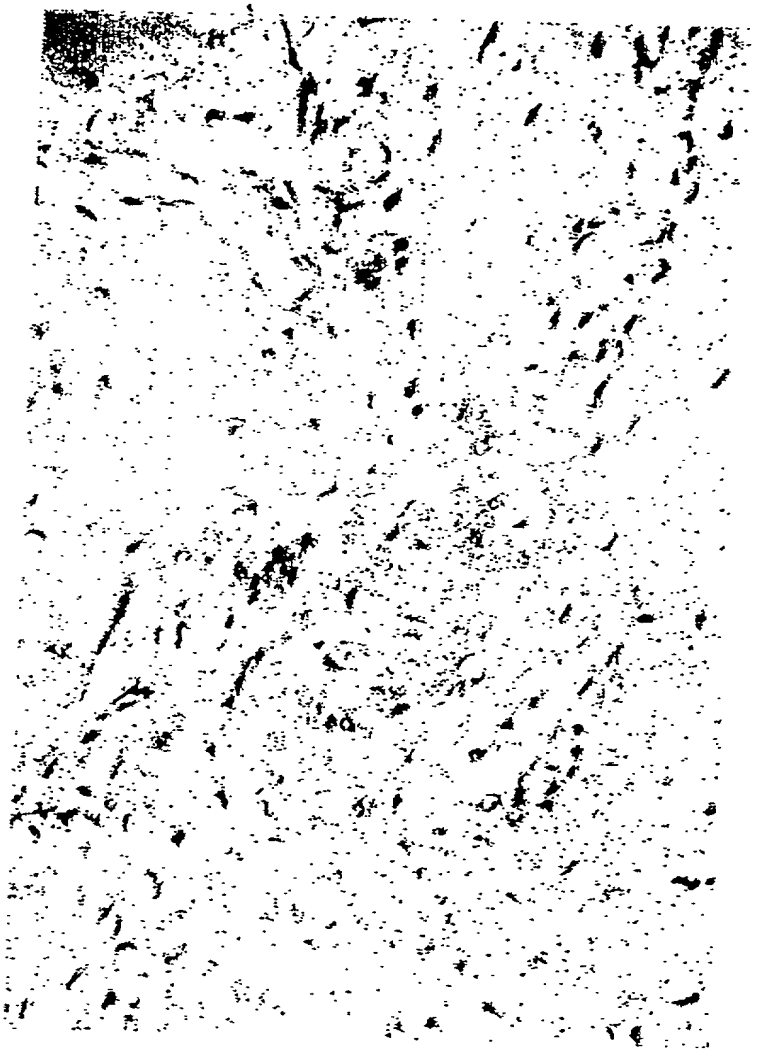


Figure 34

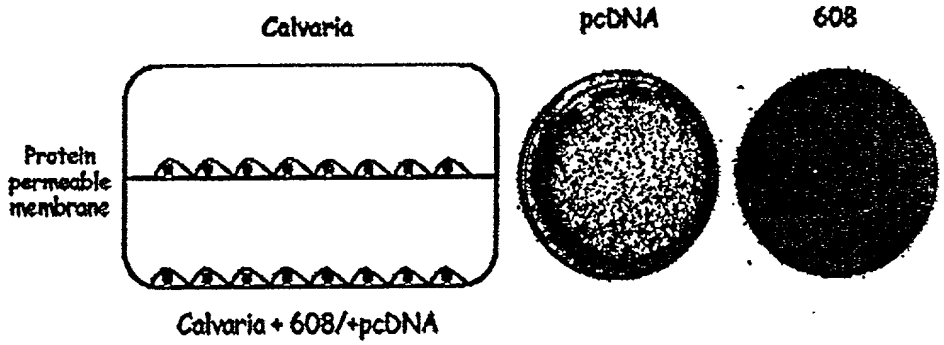


Figure 35

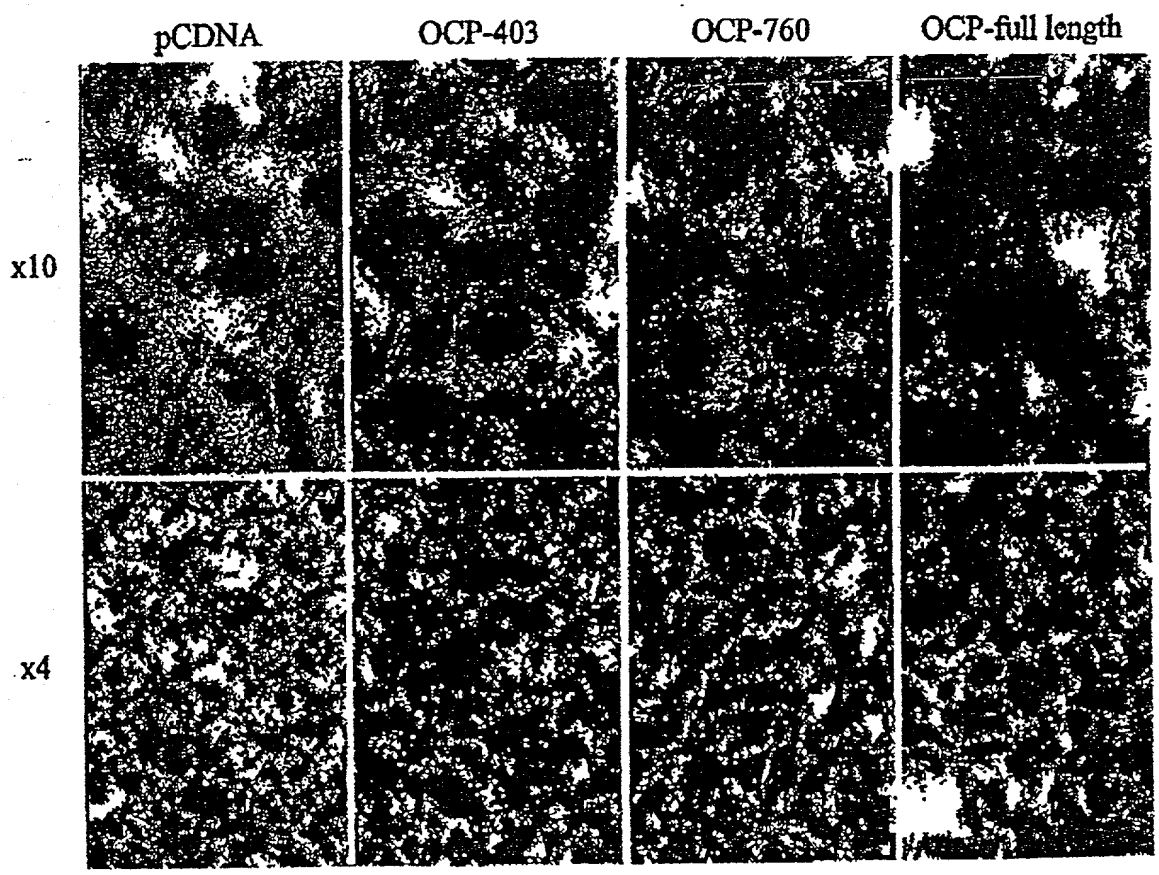


Figure 36

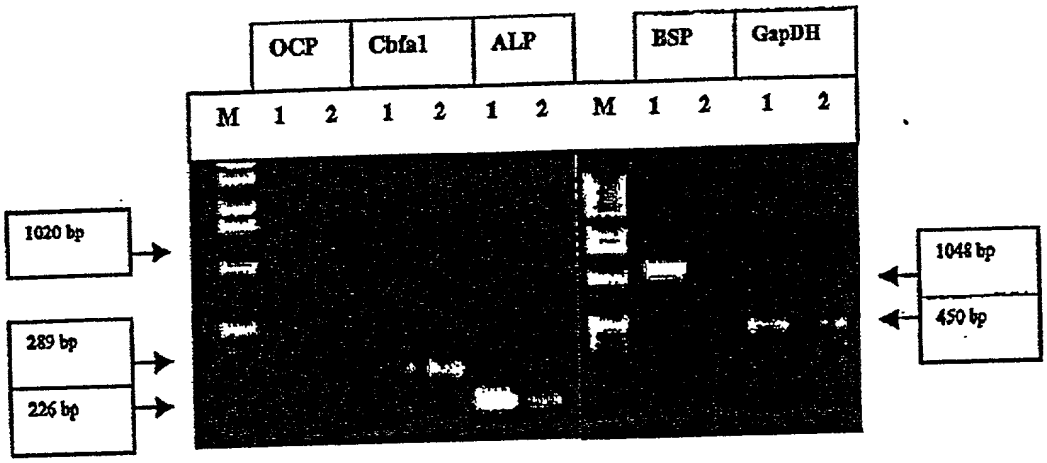


Figure 37

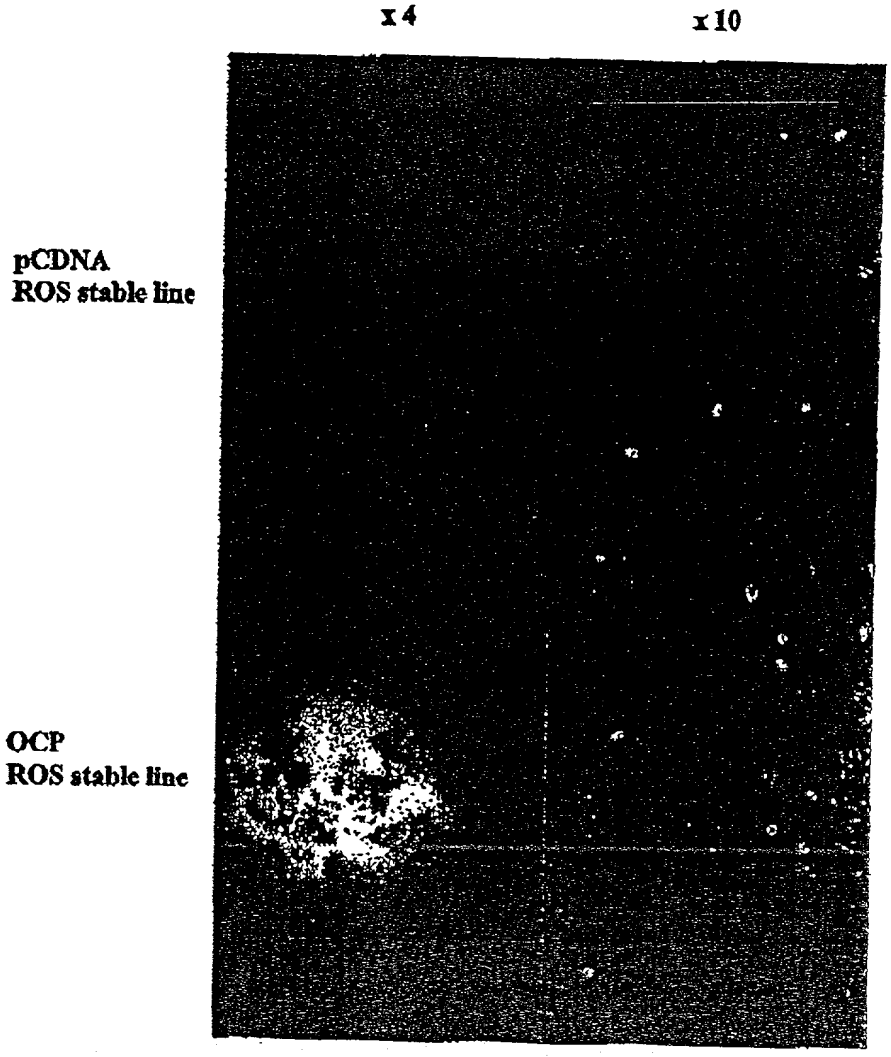


Figure 38

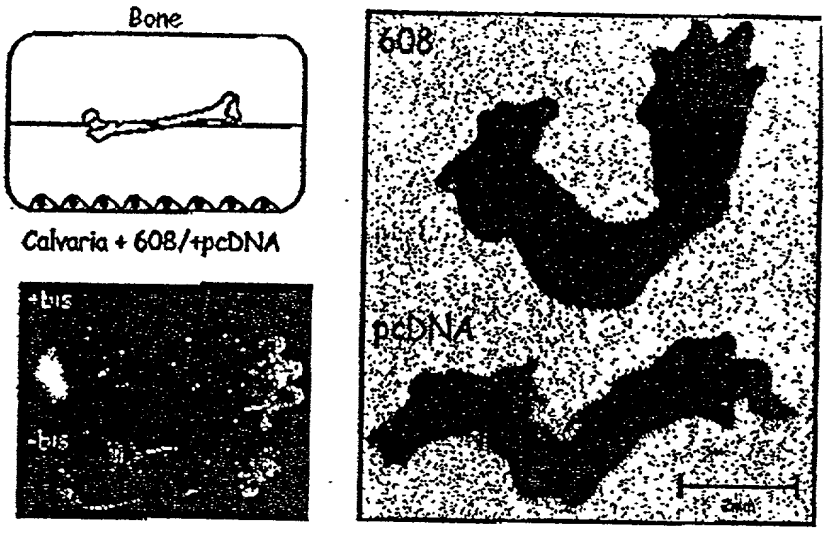


Figure 39

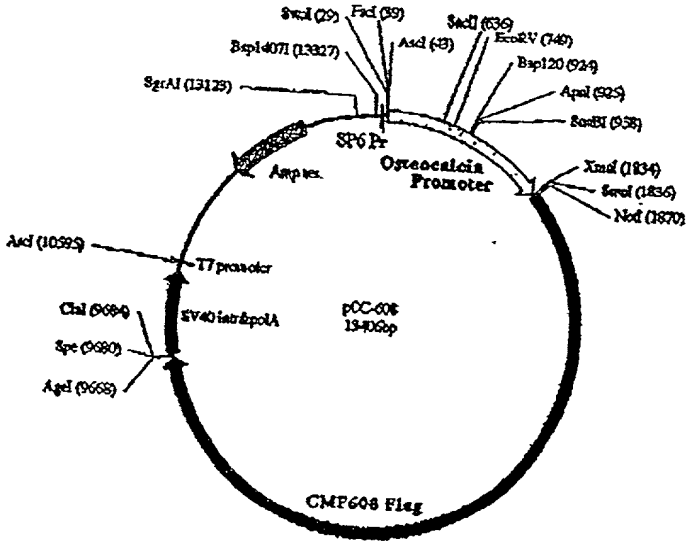


Figure 40

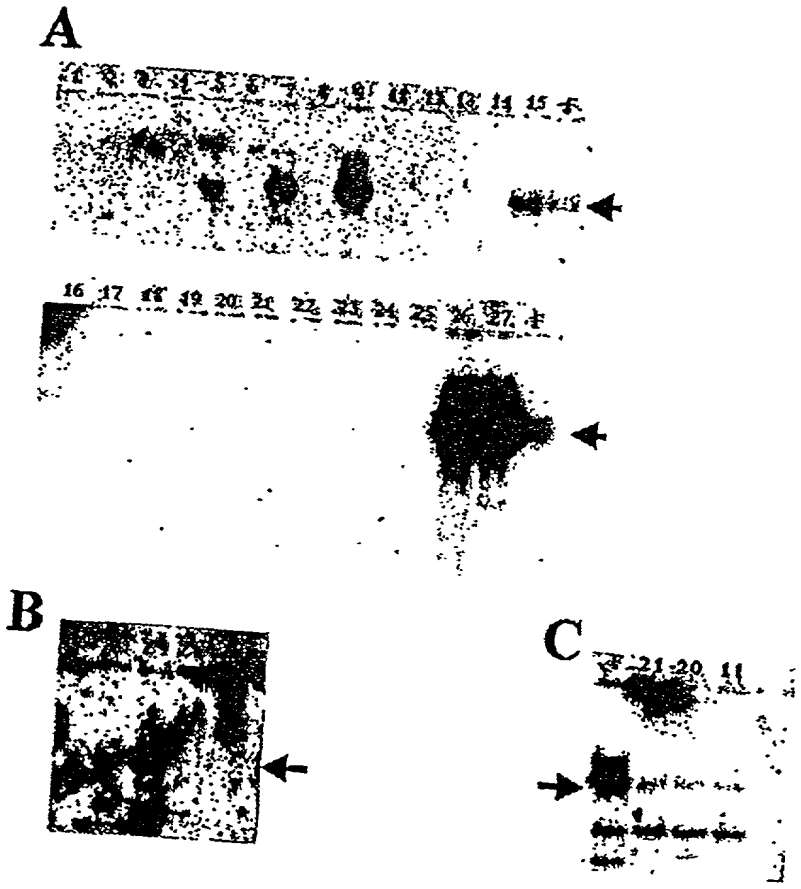
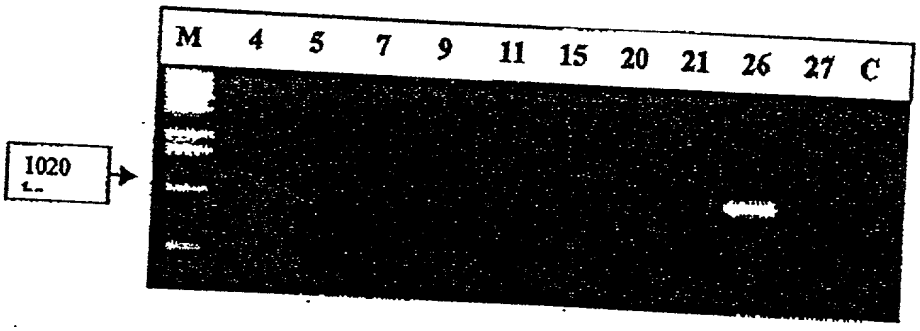


Figure 41

A



B

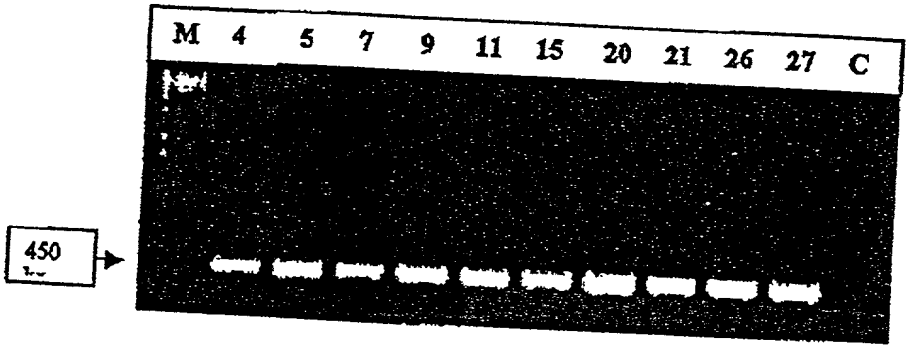
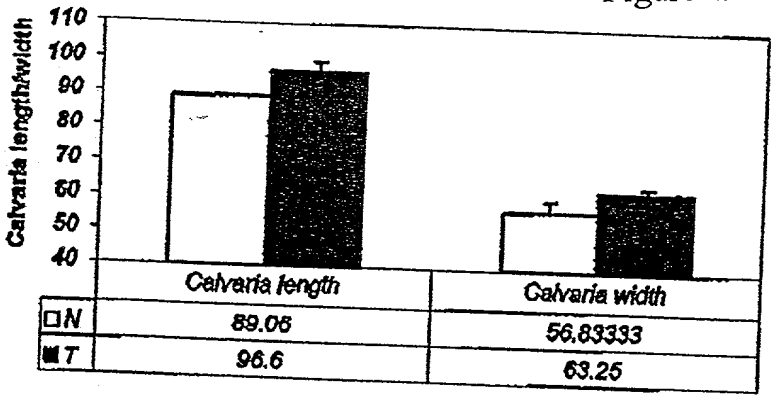


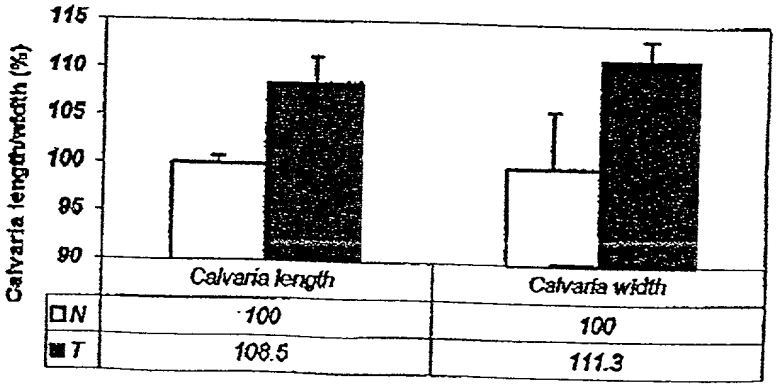
Figure 42

A



	Calvaria length	Calvaria width
N	89.06	56.83333
T	96.6	63.25

B



	Calvaria length	Calvaria width
N	100	100
T	108.5	111.3

Figure 43

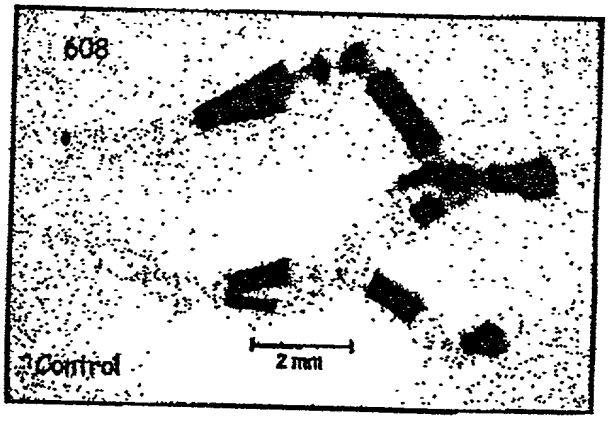


Figure 44

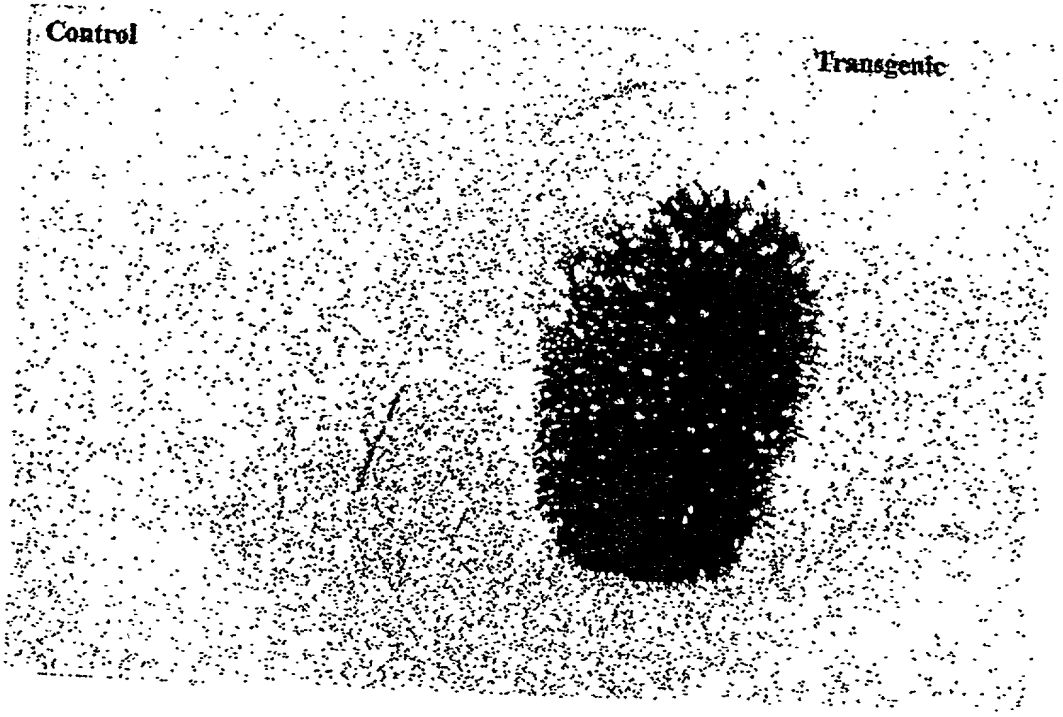
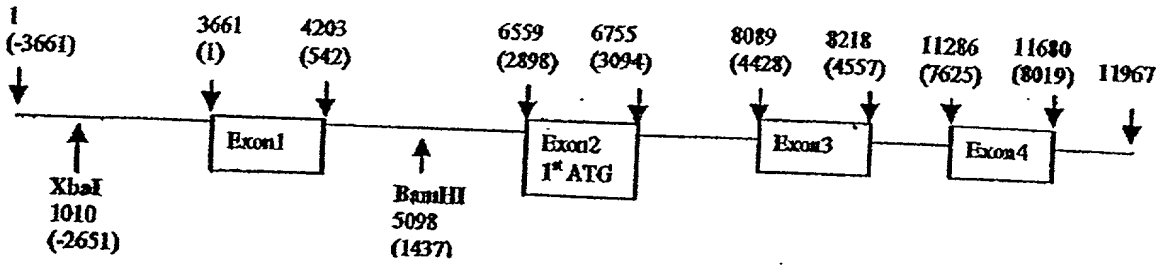


Figure 45



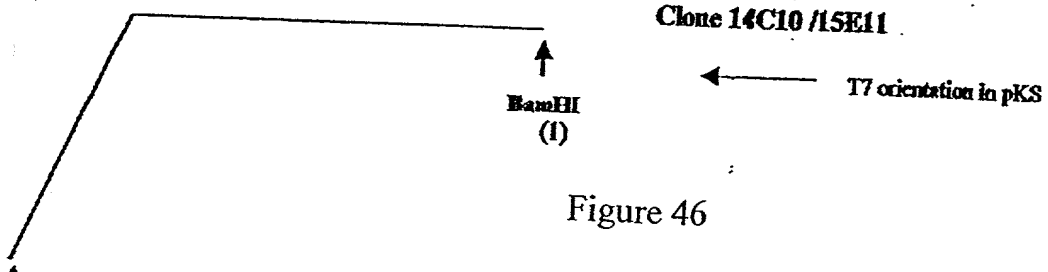


Figure 46

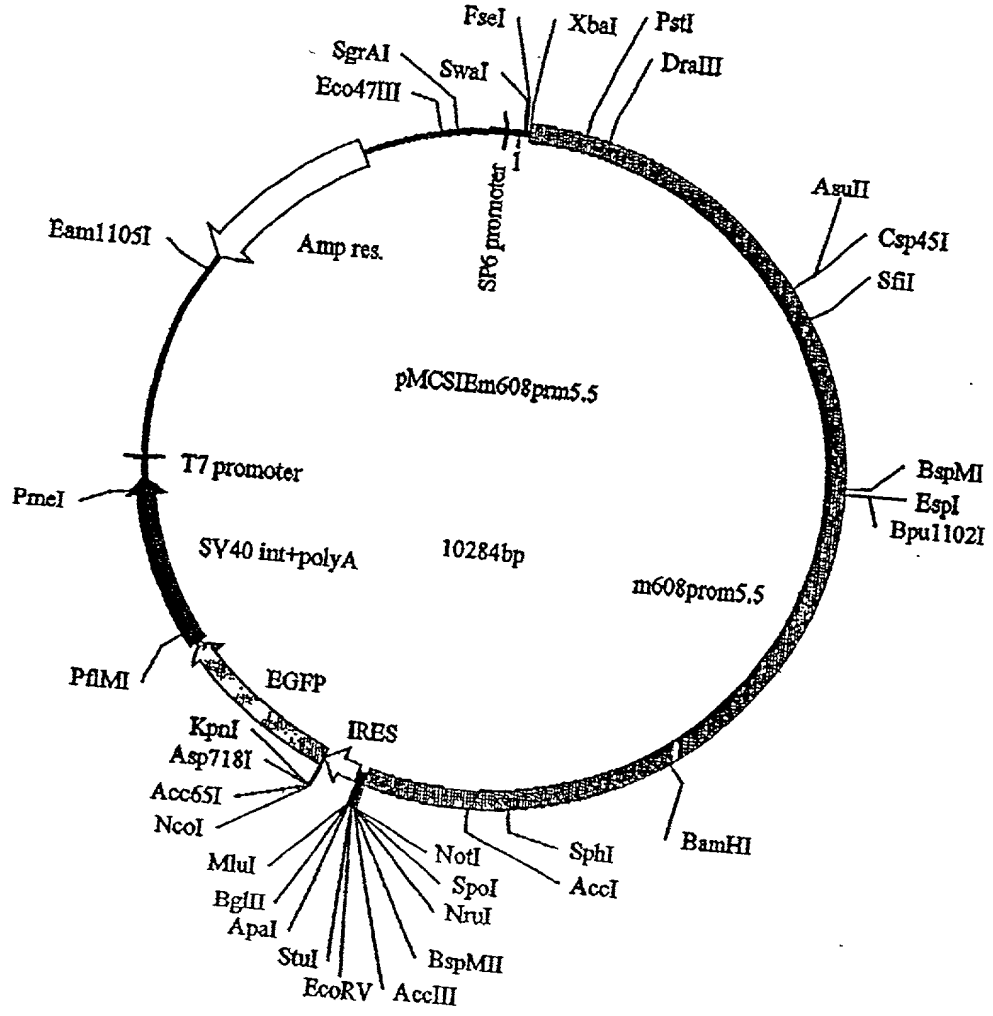
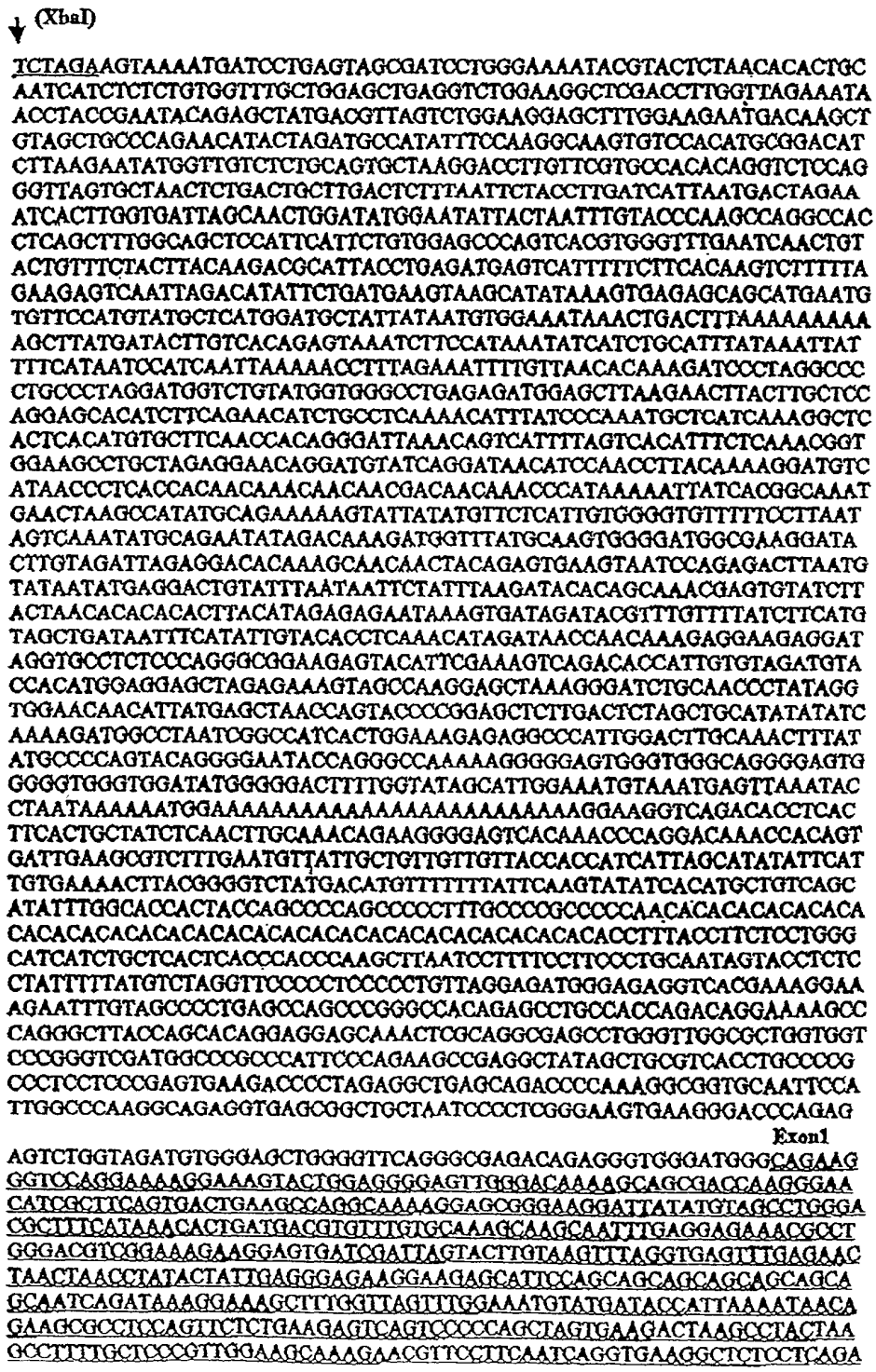


Figure 47



AGATTTCCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAGG
TATTGCCAACTCTTTTGTTAAATACAGTTTGAGGCTTAAGTGTACGGGAACTCATGTGGTA
TTCATTTACGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCATATAGTCTCTTCTGTTTCC
CAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTTAATCATC
ACGCAGAGAGTATTCAAAAATATTCAGTGATGTAACAGTGACAGTGTAGGCATAGAAGTA
ATCATTAGTAAATCTTAATTTGGGTTAAACTCATTCATAACAGCTCCAGGTTGGGAGGGAT
CACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTTTCTAACAAGAGAAGCAGAATTCTTC
CTTGGCCATGCTCCCCATCACTGTGTCAGTAAGCAGAGGGGTGTTTCCAAGCAGAGAAAG
AGCAGACAGTGTTATGCCTGCAAAGTCAGAGACTCAGCCCTCCCAGCTGGTCAGTTTACT
GTCCTCCCGGTCATTAGTTGGCTCTGAAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGC
AGACATGCTAGAAAGAAATTTGACCTTTTTTTCTAGTGGGTTATTACAGCTGTAAAAGTAT
TTTGGAAGGTTAAGCCAAATAAATAAAACACATATTAAATAATACAATGTTACAAAAATT
GATCATATAAAGAAGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTTATCTAT
TTACTGGTGCAAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAATATTTTT
AAAATGTATATAAAAGTGTAAAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAAACT
(BamHI)
TTAATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCAATG
GGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAAGTTTACA
CTCTAGGGAATAAGAGTTCCTCTTTCTCCAGATTTCACTAGCATTTGTTGTCATCATTTATC
TTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAGGAATGTCAATTTGG
ATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCACTAGTCTATTCATTTATTGGATAA
TTGGGGGATGGTGTTAATTTTTTTGCAGTTCTTATGGAATTCCAAAAAAACAAAAAACAAAC
AAACAAACAAAAAACCTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAG
AGAAATCTGCACAGAATTTATTGCTACATTGTTCATTATTCACGACAGCCAAGAATGTGGA
ACCAACTTACGTAGCCGTCAAAATATGAACGGATAAAGAAAATGTGGAAATGTGTACAAC
AGAGTCCCATGTGGCCATAAAAGAGTGAAATCATGACATATGCAGGAAATGGATGCAACT
GGAAATCAATTGGGCTAATCAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTT
CTCTGACAAACAGAAGCTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATT
TAGTTATACACTATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACA
AGCTAGCTGGGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA
GTAACATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGA
GGCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAAAGT
TTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAGGGGCGTTCCAGTTTT
CTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGCATGATGGATGG
TTTTTGTGCTATTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGTGGTGGTTGTTGTGGTGGTG
GCTGTGGTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAACATGAAATTGGGTGGGTAGGA
AGCTGGAAACGATCTGGAAGAAGTTGGGGAAAGAGAAAAATTGTATGGAGCATATTTAA
ACAAACAAACAAACAAACAAAAGGTTCATTTTGCCACAAAAAGGTGTGAATTAAATTAAC
CAGTTACGACTCTTAAAGAAAATATTCCCAATTATTCCCAGAGTTGCTATGTATGCTGTGC
CTAGGACTTTGCTTGAACTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGG – Artificial
NotI site

(SEQ ID NO:17)

Figure 48

CACAGACCTTCCTCTTCTAACCTCTCTCCCCCATCTTGTTGCTTCATCCCAGACTTCAACAC
CAGCAAGCACACTCTGCTAATGCAAGGGCTGCTCCTGTCAGGACAACAAGGAGGCTGAAG
GCAGACCCACACGTTTCCAACTGCTCCTGAGAGTCAATCCCCCTAGACTCATCTATAGCAG
GAAACCTGCTGTGATCTCCATTTCTTCTCTGACCACATCCCCAAGTTATCACAAGGAGTTT
TTCCTCAAACCTTTCCTCTCCAGCAAACCCCTTCAGCTCCTTGGGTACTTTCTCTAGCCCCT
TCATTGGGAACCCTGTGCTCCATCCAATGGATGGCTGTGAGCATCCACTTCTGTATAGAAT
CTTGGTCAGTGCAGTCTTTGTATCCTCAAGAACACTGGGTCTGAAAATTTTAACCCAAAG
AACTGTTTTTTGTTATGATTGCTGCAATCTCTTTCAATTCCAATAAAGAGTAAGCATCTCAT
TCCTTTGTCTCCTCCTTTCAGTACCACCCTGCCTTTGCTGCCTTTCTCAAAGAATCAATAAA
ACCAAAGTGATATAGATTCATGGCATTCCTCTAACTGCTACATCCACTCCAGTAGTATCTC
ACTTGGCAGGTGTAAAAGCCTGGAAGCAGTCACGAGGCAGTTTCACAGAAACTTAGCCTC
CTGGAACCTTGGCATTCCCATAGCTAGAATGCCCCAGATTTGTCCCTGAGATATTGTGGTG
GGTCTTGCATGCTTTCTTGCAGTATTTTACTGGATAAGAGTTAGAAATCTCAGGGCGAGCT
TAGCAAAAGTATACCTAGAATCTTCATGACAGTCAGGTATTGCAAACTACATTGCATATTA
GAAGAAAGTTGGTAAATTCTTCTGACAAATGGAGATTCCCTACAGATAACTTAAAAGAAC
AGCTAAGTCACACTCATATGCAAGAATTTACCAAGGCCTAGGAAAGGGGGGGGGGGTACT
GCTTTATTCATGATAAGGTCTGCTAGAGCAGAACCCCCTGGTGCTAGCTTTCACAAGGTTC
AAAGGTGTAGCATAAATTGTGACTAGAGTGTGAAATCTTTACCTGTCATTAGCTGACTCTA
GGCAGAGCTGTTTTATCTTTACTGTAAACATTACCTGGTTCCTGTCAGTCCTTTGAAGGCAT
TCCTCTGTTTTGTGACAGATACTTCTATGTACCTCGCCTGCTGTGACACCCTACTCCTTTGT
TTTCTGTATTATATAAGCCTGGTGTTCCCTTTGTGAAAAATTACATCCAGATACAGCACTC
CCTTGTGTCTGTGTCCTTTTGTCATTTCTGGCCAACTCCATGCCCACCTGCCAGAACCCCTA
GTCTTTTCCACAGATTGAGGGAGGCCGACTGAGCCTGGTCCATGGCATCTAACCACTGTCA
GCTCACTGTTGGTGACTACCTCAAGGTACAAGCTCCATTACTAATGAAACAAAATTAGAT
AAGTGTGGGTCCAGGAAGCAGGTTGTACACCCTGTCTGAATGAACATTATGAAATGACTG
AAATAAGTTAACCCATCTCTTCCTCGTTTGCTAATATAGCAAATAAACCGAGTTTCTGAGC
TGCTGCTGGTGTGTCTCCATCAGAGGGCAGAGCCAGTCTGATCCTAGCTTTCCTGTATGTG
TGTCCATTGTTTCTTCAGTTCCTGTTGCCCCATTAGGAAATCCTAAGCCATGAAAGCCATG
AATCTGGGAATGACTTTTCTAAGAAATGCCACGTGAACCTTGCGTTTCAACGTTTTGCCTG
TAAACAAGATATATGGTGCGCAGTTTATAATCATAATAAGCTTTGAAATAATATATAACTC
CATTCTCATTCTGCTTCCACGCTGAGCATCCTGTTTCCCCAGGGACCACAAGAGCATTTGA
AAAGTAGTGATTTATGACCTGCTTTGTTCTGTTACTATAAAAGCTTCATGAAAGGGCAGCC
ATGTTGAAACATGGAACTTGGGGTGACCTGTATCTGTGTTCCTGGGTCGTGCTCACTCATA
TTTGTCTCCAGAATAAATGAGTTTATCAACTTCGAGGAAAAAGTTGTGTGTTTGTATAGCA
CGCCCGTGGAGTCCCACCATTCTACTTCCTGTAATCTGTATATGGTAGAAAAAGTTAATTT
ATGTGATTCTTCCAACTCCAAATATTTCAAATCTTTTAGCCCCTCAGCCTGGGATTTCTTTG
ACTAAGTCTATTGATTTGGAAGATCTCAGTGGTTAGGATTTGCAGTCATGATGTTCATACG
TCAGGCTAAGCTGAAAAATATGACAAATGAAATGTCAAATGTCATGTGCCTGGGAATGTG
AGTGTTAGGGGGTTTTAAAGAAACAAATACCTACTCTAAATAGTTAATAAGTCCCATGGTT
CTATTCTAGTTTTGAATAATGTTCCCTAGTATACAGCAATTTAATTTGAAATGAATAGCTTC
TTATCTTGACCAATCTCAGTGACTTCATCCGTCCCAAGTCATGTTTTCATATTCATAAGGAT
AGGTCTCATTCAACCACATGTTTATCATTTGGGATCTGCATTTTTCTGATGCAAAATGATTT
ATTCTTCCAGAGCACTGGAATTGGGTTGAATCATCTTATAACGGCCAAAACTAAATGCTTC
TGTGCTAAACAGAGAGTTACAAGACCTTTTTATGTGGATGGCAGCATTTTAGTCATCCTTA
TGACAGAATGTCAGAGTGGAGCTCCCACTGGGGGAGGGGCTGGTCCTTGGCAGGATTCTC
TTGGAACATCACACAAAGAAATTCCAAATTATGAAATGCACATGATCCATCCAGAATGTG
ACTTTTGACTCTTGAACATGAGCTTTTAAAGTACGTTTGGCTGTTCAGACCTTGACTTTGAG
GTGAAGGAAAGCTCGCCAACTCCTTTTTATATGTAACACAATATATCAAGATCTAATGTGA
GACAGTATGCCAGTCCCAAGATCTGTCAATATGACTGAAGACACATTGCGATGTTATCACT
AAGGCAGGAGAAGGCAAGCTACAGTGAAGCCCAGTTCACTATAAAGCTTTATGAGAAATT
AGATAAGAAGGGTTTCTAATTTTTAAATTTTTTTTATTAGATATTTTCTTCATTTACATTTC
AAATGCTATCCCAAAAGTCCCCTATACTCCCCCCTCACCCTGCTCCCCTACCCACCCACTC
CCATTTCTTGGCCCTGGTTACTTGTGATAGTGGTCATATGATCCACCAAGCTTTACATGCTC
ACTATCTGGTCTATTGCAAGAATGGCTGCCGAGCTGATGCAGTCAGATACAGACACCTAC
AGCCAAACAGTGGAAGGAACTTGGGGACTCTTATGGAAGAAAAGGAGGAAGGGTTATGG

GCCCCGGATGGGGAAAGGAACTCCACAGGAAGACCAACATACTTGGTCAACTAACCTGG
ACCCTTGGGGCTCTCAGAGTCTGAACCACCAACCATAGAACATTCATGGGCTGTACCCAG
GCCTCTCCACTCATATGTAACAGATATGTGGCTTGGCCTTCATCTGGGTCCTGAACAACTA
GATGGGGGTTAGGGGTGGGGATGGGGGTTATCTCAAAAGCTGTTGCCTGTATGTGGGATA
TGTTCTTACTGAGCTGCCTAGTCTGGCCTCAGTGGGAGAGGAAGCACCTAGCTTTAGCCTT
GTAAAGACTTGAAGTTCTGAGGTGTCGGTGGAGGGTATACTCAGGGAGGCCCTCACCTGC
TAGGAAGAGAAGAGGAGGGGGAAGACTTGGGGGAGGGGGCAGTGAGCAGGTTGGTAAA
ATGAATAAGAAAAAAAAAATAAAAATAATTAAAAAAAAAAAAGAATGGCTGCTGAGCCC
TACTCTAAAACCATTGCATCCCCCCCCCCCAATCATTCAGTGACTACGAATTAAAATCATT
GATACTAACAATAGATGTAGGAAACTATTGTTAACTTCTTTGTGACCACGAGTGGTATTTG
GAACCTTTTTTATTGAAGCTTTCACACAGAGCCTTGTTCTTTCATTTCCCTGTACATGCATG
TAGCTTAATGATGTTCAGTGAATTAAAAATAACAATGAAGAATAAAGACAACTGTATTTT
AAGGATTCTTCGTATATATTTAAAAATCTAAGGTGGTCACCTGGAAGAAATGTCTTCAGTT
TTTCTATATATGTTTACTCTATCGTATGTTAATTAATTATATGCAATAATTCATAAAATCTA
CAACATAGTATGTAACTTATAAGAAAGTAAAACATTCATGAAATTGTGAAGGTTACTTTTC
CTTACCCTCAGAAACACTGGGTTTGAATAATTCTTATTTTGGTATCAGTGAAGAATTTGAA
AGAATGTAATAACCTACTAAGGCAAACATAGAAGTTGAAATTAAAAAGAGTAGACAGGA
GAAGTAATAAGGCAAATAATGAATATTTGCTTTAAATAGTTCTTAATGTATCATCTAACTA
GGGTGTGATTCTCCAGACTTGACTCCATCCAAAATATCCAAAATGACTCTAACCACAGTCA
TTGAAACAATGTGTTGAAAATAATAAACATTTCCTACTTGAAAATTCAGATTTCTCCTACT
TTGCTTTTTATTGCTGTGATAAGCACCATGACCAAAGCAGCTTAT

(SEQ ID NO:18)

Figure 49

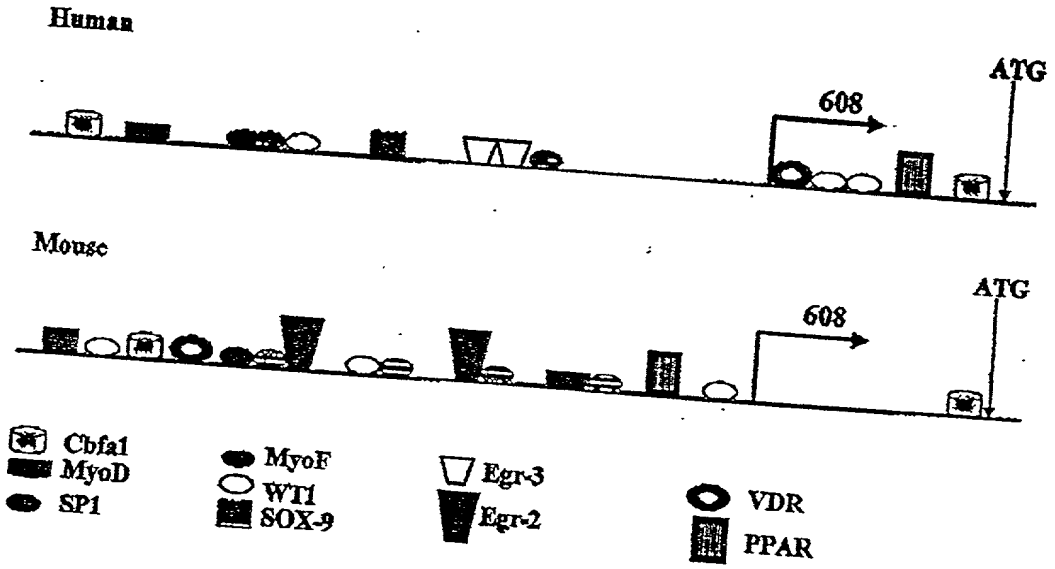

Human
ATG
608
Mouse
ATG
608
Cbfa1
MyoD
SP1
MyoF
WT1
SOX-9
Egr-3
Egr-2
VDR
PPAR

Figure 50

TAAGCCTTTTGCTCCCGTTGGAAGCAAAGAACGTTCCTTCAATCAGGTGAAGGCTCT
CCTCAGAAGATTTCATGTCTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA
AGAGCTCAGGTATAGCCAACTCTTTTGTTAAATACAGTATGAGGCTTAAGTGTACGG
CAACTCATGTGGTATTCATTTGCGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCA
TATAGTCTCTTCTGTTTCCCAGCTACCTTGCACCATCTTTGTTTATCTAATAATAGCA
AGCTCATCTGCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTA
ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTTAAACT
CATTCATAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGGAACCAACCCAGATGCCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA
TTTATCCAATGGAATACTACTCAACTTATTAAAAACAACGACTTTCATAAAATTTTTA
GGCAAATGNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACAC
TCATGGTATGCACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAAGGGA
AGAAGACACATAGCTTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGT
TTGGCTTATGAGTGAATCTTGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAC
TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA
CCACAGTCCAGGAAACTGAGTCTTAAGGATTCATATTAACTCCAGTTCCCCATGTGC
AGCTCTGAGACTTTGGCAGATCAGACACTTAACTTCACCAGCTTCCTACACAGAGCA
GTTACTATCCTTGCCTTCACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT
GACTTGTTAATAATCTTTAAAATCCAATTGTGTGTAAAGTATGTGACCAAAGAGCAT
GGTCATGCTATTAACCTTTGATGTTCTATGGACTCTTAATTTTATGGTAGAAATGTCA
ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCCTTTCAGTTTTG
AAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCCTAAAGCAATTTAACCCCCCAA
GTAGGCATAATTTTAATGCTTACTTCATCAGAATATATCTAATTGACTCTTCTAAAAA
GACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGACATAATAAATAGG
AGAAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTTATAG
GTTATATAAGAGGTGGGAAAGGAATATTTGAAACATCTAGAAGTAAAATGATCCTG
AGTAGCGATCCTGGGAAAATACGTACTCTAACACACTGCAATCATCTCTCTGTGGTT
TGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATAC
AGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCC
CAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAG
AATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTCGTGCCACACAGGTCTCCAGGG
TTAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTCCCTTGATCATTAATGACTAGA
AATCACTTGGTGATTAGCAACTGGATATGGAATATTACTTAATTTGTACCCAAGCCA
GGCCACCTCAGCTTTGGCAGCTCCATTCATTCTGTGGAGCCCAGTCACGTGGGTTTG
AATCAACTGTACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTTCTT
CACAAGTCTTTTTAGAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

TGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAAA
TAAACTGACTTTAAAAAAAAAAGCTTATGATACTTGTCACAGAGTAAATCTTCCATA
AATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTAAAAACCTTTAGAAA
TTTTGTTAACACAAAGATCCCTAGGCCCCTGCCCTAGGATGGTCTGTATGGTGGGCC
TGAGAGATGGAGCTTAAGAACTTACTTGCTCCAGGAGCACATCTTCAGAACATCTGC
CTCAAAACATTTATCCCAAATGCTCATCAAAGGCTCACTCACATGTGCTTCAACCAC
AGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGTGGAAGCCTGCTAGAGGA
ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCATAACCCTCACCA
CAACAAACAACAACGACAACAAACCCATAAAAATTATCACGGCAAATGAACTAAGC
CATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTTCCTTAATAGTCAA
ATATGCAGAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTT
GTAGATTAGAGGACACAAAGCAACAACTACAGAGTGAAGTAATCCAGAGACTTAAT
GTATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGT
ATCTTACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGTTTT
ATCTTCATGTAGCTGATAATTTCATATTGTACACCTCAAACATAGATAACCAACAAA
GAGGAAGAGGATAGGTGCCTCTCCCAGGGCGGAAGAGTACATTCGAAAGTCAGACA
CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG
GGATCTGCAACCCTATAGGTGGAACAACATTATGAGCTAACCAGTACCCCGGAGCT
CTTGACTCTAGCTGCATATATATCAAAAGATGGCCTAATCGGCCATCACTGGAAAGA
GAGGCCATTGGACTTGCAAACTTTATATGCCCCAGTACAGGGGAATACCAGGGCCA
AAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGGGGACTTTT
GGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGGAAAAAAAA
AAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTTG
CAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTTG
AATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCATTGTGAAAACTTA
CGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGC
ACCACTACCAGCCCCAGCCCCCTTTGCCCCGCCCCCAACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCA
TCATCTGCTCACTCACCCACCCAAGCTTAATCCTTTTCCTTCCCTGCAATAGTACCTC
TCCTATTTTTATGTCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTCACGA
AAGAAAGGAATTTGTAGCCCTTGAGCCAGCCCGGGCCACAGAGCCTGCCACCAGAC
AGGAAAAGCCCAGGGCTTACCAGCACAGGAGGAGCAAACTCGCAGGCGAGCCTGG
GTTGGCGCTGGTGGTCCCGGGTCGATGGCCCGCCCATTCCCAGAAGCCGAGGCTATA
GCTGCGTCACCTGCCCCGCCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC
CCCAAAGGCGGTGCAATTCCATTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCC
CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTTCAGGG
CGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGG
AGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAAGCCAGGCA
AAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGATGACG
TGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAGGA

GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTTGAGAACTAACTAACCTATACTA
TTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAGCAATCAGATAA
AGGAAAGCTTTGGTTAGTTTGGAAATGTATGATACCATTAAAATAACAGAAGCGCCT
CCAGTTCTCTGAAGAGTCAGTCCCCCAGCTAGTGAAGACTAAGCCTACTAAGCCTTT
TGCTCCCGTTGGAAGCAAAGAACGTTCCTTCAATCAGGTGAAGGCTCTCCTCAGAAG
ATTTCCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAG
GTATTGCCAACTCTTTTGTTAAATACAGTTTGAGGCTTAAGTGTACGGGAACTCATG
TGGTATTCATTTACGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCATATAGTCTC
TTCTGTTTCCCAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCT
GCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTAACAGTGACA
GTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATTTGGGTTAAACTCATTCATAA
CAGCTCCAGGTTGGGAGGGATCACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTT
TCTAACAAGAGAAGCAGAATTCTTCCTTGGCCATGCTCCCCATCACTGTGTCAGTAA
GCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGTTATGCCTGCAAAGTCA
GAGACTCAGCCCTCCCAGCTGGTCAGTTTACTGTCCTCCCGGTCATTAGTTGGCTCTG
AAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGCAGACATGCTAGAAAGAAATTT
GACCTTTTTTTCTAGTGGGTTATTACAGCTGTAAAAGTATTTTGGAAGGTTAAGCCA
AATAAATAAAACACATATTAAATAATACAATGTTACAAAAATTGATCATATAAAGA
AGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTTATCTATTTACTGGTGC
AAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAATATTTTTAAAATG
TATATAAAAGTGTAAAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAAACTTTA
ATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCA
ATGGGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAA
GTTTACACTCTAGGGAATAAGAGTTCCTCTTTCTCCAGATTTCACTAGCATTTGTTGT
CATCATTTATCTTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAG
GAATGTCAATTTGGATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCACTAGTC
TATTCATTTATTGGATAATTGGGGGGATGGTGGTAATTTTTTTGCAGTTCTTATGGAA
TTCCAAAAAAACAAAAAAACAAACCAACCAACCAAAAAACCTCTGAAACTAGAACTACC
AATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACAGAATTTATTGCTACATTG
TTCATTATTCACGACAGCCAAGAATGTGGAACCAACTTACGTAGCCGTCAAAATATG
AACGGATAAAGAAAATGTGGAAATGTGTACAACAGAGTCCCATGTGGCCATAAAAG
AGTGAAATCATGACATATGCAGGAAATGGATGCAACTGGAAATCAATTGGGCTAAT
CAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTTCTCTGACAAACAGAAG
CTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATTTAGTTATACACTA
TTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTG
GGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAAC
ATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGAG
GCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAA
AGTTTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAAGGGGCGTT

CCAGTTTTCTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGC
ATGATGGATGGTTTTTGTGCTATTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGTGGT
GGTTGTTGTGGTGGTGGCTGTGGTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAAC
ATGAAATTGGGTGGGTAGGAAGCTGGAAACGATCTGGAAGAAGTTGGGGAAAGAG
AAAAATTGTATGGAGCATATTTAAACAAACAAACAAACAAACAAAAGGTTCATTTT
GCCACAAAAAGGTGTGAATTAAATTAACCAGTTACGACTCTTAAAGAAAATATTCCC
AATTATTCCCAGAGTTGCTATGTATGCTGTGCCTAGGACTTTGCTTGAACTGGCCCTA
TAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCAGGGAAGTCAGCTGCT
TGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGGTCTGTC
CTCGCCGATGTGCCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCGGGACCTGA
CCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGTG
GACCTTGCCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCCTGGTACTTTGCCCC
CCAAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTCATACAAACCTT
ATATTGTGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACTTGAAGCC
CCTCTTAGTTGTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTT
AAATGAAAAAGATACAGTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAAT
TGTCATTGAATATTTTAAAAAAGCATTTCTTTTGCAATTTATAAATACCCATTACAAA
ATGGCTTACTTAAAATACTTGCCTTACTAAATCTGACAAATTATGGTGATATTTTGAA
GGTTTATGAAAATTTGTTTATGTGTATAAATGCACAAGAAATGGGATATGCCATCAC
CTATGTGCCATTAGTGAGCATGTACAGTATGCCAAACACTATTGTTCACGTTTGGAG
GAAGTAATGGGGGTGGGGGAGCAACAAGGGTTATAACCGTATACCCAGTGCCTTGG
AAGCGATTGCAAACAGTAAAGACTGACATTGTGTTCTCCCTATGAGGGAGGGGCCTT
GGGCTGAGCACTTTGCAATGAGCATTTGCTCATTGTGCTGGCAGGTTTTATGATAAC
TTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACTTCAACTGAGAACATGCCTGA
AGAAGATCAGATTATAGGCAGGCCTGTGGGGCATTTTCTTAATTAGTGATTCATGGG
GCAGGGCCCAGTCCATTGTTCGTGGTACCATTTCTCAGGCACTATTAAAAAAAAAAA
AACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGATCTC
TGCATCAGCTCCTGCCTCCAGGTTCCTACCCTATTTGAGTTCCTGTCCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAACTT
GCTGTTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGT
TTATAACACCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAATGGAGCATAGA
AAGTCTCCAGACTACACCAGAGTGTAATCTGGGCCTGAGCTTAGAATCACACCCAC
GTGCACTCCACTGCCGGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCT
GTTTGTTTCCATTTTGACATGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATG
ATTTTTTTTCCAGGTATAACAGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCT
GAGCAGACTGGAGTTACTCATGCTGCACAGCAATGGCATTCACAGAGTCAGTGACA
AGACCTTCTCGGGCTTGCAGTCCTTGCAGGTGAGATAGGTAGAGGGTGATGGAGGC
TGAGAAGAGAGGTGCAACTGTGGGTTATACCCAAAAGCTGCTGATTCCCGTGGGAG
ACATTCTATAAGCATTCTATAAACTAGAGGCAGATATCAAGGAAGGATTTCAATTGT
AATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATGCTGGGGAAAATGCTTAC

ACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAAAAAGCACAGCATG
GCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTCAGAATCCCAGAGACT
CATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAAACTTGAC
TCAGAAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT
GCATGAATAAGGGATCCCAGAGAGAAGGGAAGAAAAAGGAAGGAAGGAAGGAAG
GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAGAGGGAGGAAAGGAGGG
AGGGAAGGAAGGAAAGGGAAAGGAAAAAAGAGATGGGGAGGGAGGGAAGGAAAG
GAAAGGGGGAGAAAGAAGAGAAGAAAGGAAAATAAATAAATTTTCAGGGATTATT
ACACCTTTAAATTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGT
GGGATCCCTTATATAAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAA
ATTTTGAGTTTTCTCTACTTTTTATCCTCTACCATCTTCAAACTGAAACTACAATTATT
CCCACAAATGAAGAAAATGCTGTAAGAGTTTTCACACACCGAAGTGGGAAACTTAA
GGATTAGACAAGTCTAACAATGAGAATGGGGAGAACAAAAAGAGACTGCACAGGG
AGCCCTTTCTCTGCTTATAATCTTGACACTTGAGAAGCTAATTGACGCTGCATGACTA
CTCAACTCTTTAAGCAAACAATGCTGTTGTTCATGAAAAGCACAATAAAGTACATAT
GTCCCATAATATTCATCAAAATTTGCATGCAGCACATAATAGCAATCAAAGCAATAA
CACCCACTGTTCACAGAGACTTTAAACATGAAACTGGAACTATGTCTAGTGTTTTGA
CTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTGATTTAGGTCATCAGA
CAGCATTTGGAACATGTATCTTCAGGAGGAATCATTCATGTATCCTGCATGAAATTC
TCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATTGGGTTTG
AGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGGCTTT
TATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA
GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTTATTAGTATTTCATGTTCTTT
TCTTTTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGA
AAATAATCAGAGAGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTT
TTTGCTTCTAATTCAGAATCAATTAAATTCATATTACTATAAAAGACAGCACGCCAG
ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACA
GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTTGAAAACATTCGGCTACATG
ATGGAAGCCAGGCACAAAAAGCCACATATTGCATGGTTATGTTTATATGAAATGTTT
AAAATACATGGATTCTTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA
TTAAAAAAAAAAAAAACTATTGATGTGGAATGATCTTAATTTGGGGAAAAGACAATT
TCCTAAGACGAAATAGTTGAGGTAGATATAGTTATATCCCTGTGGATATTGTAATAA
ACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGGGCAGGAGGAAGTGAAATG
AGATGGTAGAAAGGAAAGTCATATACCATGGCTTCTCTCGTGGGTGGAATCTAGAT
ATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATCAAAACCAACA
GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAGTTCATGGTCAATGTGTGT
GGAGACACCATAATAAAACTCCTTTTTTGTTTGCTAACTAAAACCACTAAAATCTAA
AAACAAAACATTTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGG
GGAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTTAAACTGCTTTTTA
CGTTTTTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTTCACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAG
TGAGGCAATCTTGGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAG
GACAGTGAGAGGGAAATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCA
TAGTCCACTACATTACTTTGTATTCTAATATTAAGAAAATAATAAACCCATTTCTGTG
CACTTATCACCCAGGCTCAACAGTTATCTTGGCCACAGATCCTGTCTCACTGCATCCT
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCTTACTCCT
ACACAGAACTAAAGATTAAAGAGAGTTTAAAAGTAACCATGACATCTCTCTGTTCCT
TTAGCGATAAGTTCTTAATATTTATGGCTGCTTGTGTATGTTCTAATTTCTCTAATATT
GTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGTTGGAGTTAAGGTCCCATAG
GATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCTTTGTGAAAGTTCCTTT
GAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATGAGCTATAACAA
AGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGGTTGCA
CCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTT
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT
TTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTACTT
GTATGATAACTTCATTGACCTCCCTCCCAAAAGAAATGGTCTCCTCTATGCCAAACC
TAGAAAGCCTTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATTTAAAGTGGT
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACTATCTTGTTTGTTTGTTTCTTTTTTT
ATARKACGTATTTTCCTCAATTTCATTTAGAATGATATCCCAAAAGTCCCCCATAACC
TCCCCCCCCACTTCCCTACCTACCCATTCCCATTTTTTGGCCCTGGCATTCCCCTGTACT
GGGGCATATAAAGTTTGCGTGTCCAATGGACCTCTCTTTCCAGTGATGGCCAACTAG
GCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGGTACTGGTTAGTTC
ATAATGTTGTTGCACCTACAGGGTTGAA (SEQ ID NO:20)

Figure 51

MPKRAHWGALSVVLILLWGHPRVALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE
RINLGFNSIQALSETSFAGLTKLELLMIHGNEIPSIPDGALRDLSSLQVFKFSYNKLRVITG
QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSLRLLHLEGNLLHQLHPSTFSTFTFLDYFR
LSTIRHLYLAENMVRTLPASMLRNMPLLENLYLQGNPWTCDCEMRWFLEWDAKSRGIL
KCKKDKAYEGGQLCAMCFSPKKLYKHEIHKLKDMTCLKPSIESPLRQNRSRSIEEEQEQ
EEDGGSQLILEKFQLPQWSISLNMTDEHGNMVNLVCDIKKPMDVYKIHLNQTDPPDIDIN
ATVALDFECPMTRENYEKLWKLIAYYSEVPVKLHRELMLSKDPRVSYQYRQDADEEAL
YYTGVRAQILAEPEWVMQPSIDIQLNRRQSTAKKVLLSYYTQYSQTISTKDTRQARGRS
WVMIEPSGAVQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPDSKFSIL
SSGWLRIKSMEPSDSGLYQCIAQVRDEMDRMVYRVLVQSPSTQPAEKDTVTIGKNPGES
VTLPCNALAIPEAHLSWILPNRRIINDLANTSHVYMLPNGTLSIPKVQVSDSGYYRCVAV
NQQGADHFTVGITVTKKGSGLPSKRGRRPGAKALSRVREDIVEDEGGSGMGDEENTSR
RLLHPKDQEVFLKTKDDAINGDKKAKKGRRKLKLWKHSEKEPETNVAEGRRVFESRRR
INMANKQINPERWADILAKVRGKNLPKGTEVPPLIKTTSPPSLSLEVTPPFPAVSPPSASP
VQTVTSAEESSADVPLLGEEEHVLGTISSASMGLEHNHNGVILVEPEVTSTPLEEVVDDL
SEKTEEITSTEGDLKGTAAPTLISEPYEPSPTLHTLDTVYEKPTHEETATEGWSAADVGSS
PEPTSSEYEPPLDAVSLAESEPMQYFDPDLETKSQPDEDKMKEDTFAHLTPTPTIWVNDS
STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQDTLLIKKGMKEMSQTLQGGNM
LEGDPTHSRSSESEGQESKSITLPDSTLGIMSSMSPVKKPAETTVGTLLDKDTTTVTTTPR
QKVAPSSTMSTHPSRRRPNGRRRLRPNKFRHRHKQTPPTTFAPSETFSTQPTQAPDIKISS
QVESSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGTPRRKHGKRPNKHRYTPSTVSSRA
SGSKPSPSPENKHRNIVTPSSETILLPRTVSLKTEGPYDSLDYMTTTRKIYSSYPKVQETLP
VTYKPTSDGKEIKDDVATNVDKHKSDILVTGESITNAIPTSRSLVSTMGEFKEESSPVGFP
GTPTWNPSRTAQPGRLQTDIPVTTSGENLTDPPLLKELEDVDFTSEFLSSLTVSTPFHQEE
AGSSTTLSSIKVEVASSQAETTTLDQDHLETTVAILLSETRPQNHTPTAARMKEPASSSPS
TILMSLGQTTTTKPALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEGTQHMSGPNE
LSTPSSDRDAFNLSTKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPILPTA
TVRLPEMSTQSASRYFVTSQSPRHWTNKPEITTYPSGALPENKQFTTPRLSSTTIPLPLHM
SKPSIPSKFTDRRTDQFNGYSKVFGNNNIPEARNPVGKPPSPRIPHYSNGRLPFFTNKTLSF
PQLGVTRRPQIPTSPAPVMRERKVIPGSYNRIHSHSTFHLDFGPPAPPLLHTPQTTGSPSTN
LQNIPMVSSTQSSISFITSSVQSSGSFHQSSSKFFAGGPPASKFWSLGEKPQILTKSPQTVSV
TAETDTVFPCEATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG
QYMCTASNLHGLDRMVVLLSVTVQQPQILASHYQDVTVYLGDTIAMECLAKGTPAPQI
SWIFPDRRVWQTVSPVESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV
AALPPVIHQEKLENISLPPGLSIHIHCTAKAAPLPSVRWVLGDGTQIRPSQFLHGNLFVFP
NGTLYIRNLAPKDSGRYECVAANLVGSARRTVQLNVQRAAANARITGTSPRRTDVRYG
GTLKLDCSASGDPWPRILWRLPSKRMIDALFSFDSRIKVFANGTLVVKSVTDK

DAGDYLCVARNKVGDDYVVLKVDVVMKPAKIEHKEENDHKVFYGGDLKVDCVATGL
PNPEISWSLPDGSLVNSFMQSDDSGGRTKRYVVFNNGTLYFNEVGMREEGDYTCFAEN
QVGKDEMRVRVKVVTAPATIRNKTYLAVQVPYGDVVTVACEAKGEPMPKVTWLSPTN
KVIPTSSEKYQIYQDGTLLIQKAQRSDSGNYTCLVRNSAGEDRKTVWIHVNVQPPKING
NPNPITTVREIAAGGSRKLIDCKAEGIPTPRVLWAFPEGVVLPAPYYGNRITVHGNGSLDI
RSLRKSDSVQLVCMARNEGGEARLIVQLTVLEPMEKPIFHDPISEKITAMAGHTISLNCS
AAGTPTPSLVWVLPNGTDLQSGQQLQRFYHKADGMLHISGLSSVDAGAYRCVARNAA
GHTERLVSLKVGLKPEANKQYHNLVSIINGETLKLPCTPPGAGQGRFSWTLPNGMHLEG
PQTLGRVSLLDNGTLTVREASVFDRGTYVCRMETEYGPSVTSIPVIVIAYPPRITSEPTPVI
YTRPGNTVKLNCMAMGIPKADITWELPDKSHLKAGVQARLYGNRFLHPQGSLTIQHAT
QRDAGFYKCMAKNILGSDSKTTYIHVF (SEQ ID NO: 21)

Figure 52

ATGCCCAAGC GCGCGCACTG GGGGGCCCTC TCCGTGGTGC TGATCCTGCT
TTGGGGCCAT CCGCGAGTGG CGCTGGCCTG CCCGCATCCT TGTGCCTGCT
ACGTCCCCAG CGAGGTCCAC TGCACGTTCC GATCCCTGGC TTCCGTGCCC
GCTGGCATTG CTAGACACGT GGAAAGAATC AATTTGGGGT TTAATAGCAT
ACAGGCCCTG TCAGAAACCT CATTTGCAGG ACTGACCAAG TTGGAGCTAC
TTATGATTCA CGGCAATGAG ATCCCAAGCA TCCCCGATGG AGCTTTAAGA
GACCTCAGCT CTCTTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT
GATCACAGGA CAGACCCTCC AGGGTCTCTC TAACTTAATG AGGCTGCACA
TTGACCACAA CAAGATCGAG TTTATCCACC CTCAAGCTTT CAACGGCTTA
ACGTCTCTGA GGCTACTCCA TTTGGAAGGA AATCTCCTCC ACCAGCTGCA
CCCCAGCACC TTCTCCACGT TCACATTTTT GGATTATTTC AGACTCTCCA
CCATAAGGCA CCTCTACTTA GCAGAGAACA TGGTTAGAAC TCTTCCTGCC
AGCATGCTTC GGAACATGCC GCTTCTGGAG AATCTTTACT TGCAGGGAAA
TCCGTGGACC TGCGATTGTG AGATGAGATG GTTTTTGGAA TGGGATGCAA
AATCCAGAGG AATTCTGAAG TGTAAAAAGG ACAAAGCTTA TGAAGGCGGT
CAGTTGTGTG CAATGTGCTT CAGTCCAAAG AAGTTGTACA AACATGAGAT
ACACAAGCTG AAGGACATGA CTTGTCTGAA GCCTTCAATA GAGTCCCCTC
TGAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG
GATGGTGGCA GCCAGCTCAT CCTGGAGAAA TTCCAACTGC CCCAGTGGAG
CATCTCTTTG AATATGACCG ACGAGCACGG GAACATGGTG AACTTGGTCT
GTGACATCAA GAAACCAATG GATGTGTACA AGATTCACTT GAACCAAACG
GATCCTCCAG ATATTGACAT AAATGCAACA GTTGCCTTGG ACTTTGAGTG
TCCAATGACC CGAGAAAACT ATGAAAAGCT ATGGAAATTG ATAGCATACT
ACAGTGAAGT TCCCGTGAAG CTACACAGAG AGCTCATGCT CAGCAAAGAC
CCCAGAGTCA GCTACCAGTA CAGGCAGGAT GCTGATGAGG AAGCTCTTTA
CTACACAGGT GTGAGAGCCC AGATTCTTGC AGAACCAGAA TGGGTCATGC
AGCCATCCAT AGATATCCAG CTGAACCGAC GTCAGAGTAC GGCCAAGAAG
GTGCTACTTT CCTACTACAC CCAGTATTCT CAAACAATAT CCACCAAAGA
TACAAGGCAG GCTCGGGGCA GAAGCTGGGT AATGATTGAG CCTAGTGGAG
CTGTGCAAAG AGATCAGACT GTCCTGGAAG GGGGTCCATG CCAGTTGAGC
TGCAACGTGA AAGCTTCTGA GAGTCCATCT ATCTTCTGGG TGCTTCCAGA
TGGCTCCATC CTGAAAGCGC CCATGGATGA CCCAGACAGC AAGTTCTCCA
TTCTCAGCAG TGGCTGGCTG AGGATCAAGT CCATGGAGCC ATCTGACTCA
GGCTTGTACC AGTGCATTGC TCAAGTGAGG GATGAAATGG ACCGCATGGT
ATATAGGGTA CTTGTGCAGT CTCCCTCCAC TCAGCCAGCC GAGAAAGACA
CAGTGACAAT TGGCAAGAAC CCAGGGGAGT CGGTGACATT GCCTTGCAAT
GCTTTAGCAA TACCCGAAGC CCACCTTAGC TGGATTCTTC CAAACAGAAG
GATAATTAAT GATTTGGCTA ACACATCACA TGTATACATG TTGCCAAATG
GAACTCTTTC CATCCCAAAG GTCCAAGTCA GTGATAGTGG

TTACTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCAGAC CATTTTACGG
TGGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC
AGACGCCCAG GTGCAAAGGC TCTTTCCAGA GTCAGAGAAG ACATCGTGGA
GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC
TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC
ATCAATGGAG ACAAGAAAGC CAAGAAAGGG AGAAGAAAGC TGAAACTCTG
GAAGCATTCG GAAAAAGAAC CAGAGACCAA TGTTGCAGAA GGTCGCAGAG
TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG
GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA
GGGCACAGAA GTACCCCCGT TGATTAAAAC CACAAGTCCT CCATCCTTGA
GCCTAGAAGT CACACCACCT TTTCCTGCTG TTTCTCCCCC CTCAGCATCT
CCTGTGCAGA CAGTAACCAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT
ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCCTCA GCCAGCATGG
GGCTAGAACA CAACCACAAT GGAGTTATTC TTGTTGAACC TGAAGTAACA
AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA
GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA
TATCTGAGCC TTATGAACCA TCTCCTACTC TGCACACATT AGACACAGTC
TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC
AGATGTTGGA TCGTCACCAG AGCCCACATC CAGTGAGTAT GAGCCTCCAT
TGGATGCTGT CTCCTTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA
GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC
CTTTGCACAC CTTACTCCAA CCCCCACCAT CTGGGTTAAT GACTCCAGTA
CATCACAGTT ATTTGAGGAT TCTACTATAG GGGAACCAGG TGTCCCAGGC
CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTTG TGAAAAGTAG
TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT
CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC
AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA
CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG
AAACCACAGT TGGTACCCTC CTAGACAAAG ACACCACAAC AGTAACAACA
ACACCAAGGC AAAAAGTTGC TCCGTCATCC ACCATGAGCA CTCACCCTTC
TCGAAGGAGA CCCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC
ACCGGCACAA GCAAACCCCA CCCACAACTT TTGCCCCATC AGAGACTTTT
TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCAA GTCAAGTGGA
GAGTTCTCTG GTTCCTACAG CTTGGGTGGA TAACACAGTT AATACCCCCA
AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA
CGGAGAAAAC ACGGGAAGAG GCCAAACAAA CATCGATATA CCCCTTCTAC
AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA
AACATAGAAA CATTGTTACT CCCAGTTCAG AAACTATACT TTTGCCTAGA
ACTGTTTCTC TGAAAACTGA GGGCCCTTAT GATTCCTTAG ATTACATGAC
AACCACCAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC
CAGTCACATA

TAAACCCACA TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG
TTGACAAACA TAAAAGTGAC ATTTTAGTCA CTGGTGAATC AATTACTAAT
GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTTAAGGA
AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA
GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT
GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA
TTTCACTTCC GAGTTTTTGT CCTCTTTGAC AGTCTCCACA CCATTTCACC
AGGAAGAAGC TGGTTCTTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG
GCTTCAAGTC AGGCAGAAAC CACCACCCTT GATCAAGATC ATCTTGAAAC
CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCCTA
CTGCTGCCCG GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC
ATGTCTTTGG GACAAACCAC CACCACTAAG CCAGCACTTC CCAGTCCAAG
AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTTC TTGAATTATG
TGGGGAATCC AGAAACAGAA GCAACCCCAG TCAACAATGA AGGAACACAG
CATATGTCAG GGCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC
ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA TTTGGTAGTA
GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT
GCTTCTCATC AACTAACCAG AGTCCCTGCC AAACCCATCC TACCAACAGC
AACAGTGAGG CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG
TAACTTCCCA GTCACCTCGT CACTGGACCA ACAAACCGGA AATAACTACA
TATCCTTCTG GGGCTTTGCC AGAGAACAAA CAGTTTACAA CTCCAAGATT
ATCAAGTACA ACAATTCCTC TCCCATTGCA CATGTCCAAA CCCAGCATTC
CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTACTCCAAA
GTGTTTGGAA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAAGCC
TCCCAGTCCA AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA
CCAACAAGAC TCTTTCTTTT CCACAGTTGG GAGTCACCCG GAGACCCCAG
ATACCCACTT CTCCTGCCCC AGTAATGAGA GAGAGAAAAG TTATTCCAGG
TTCCTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG GACTTTGGCC
CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCCTCA
ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC
CTTTATAACA TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT
CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTCTTGGG
GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC
TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA CCAAAGCCTT
TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC
AGGATACAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA
GGTTCAAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC
ACGGCCTGGA CAGGATGGTG GTCTTGCTTT CGGTCACCGT GCAGCAACCT
CAAATCCTAG CCTCCCACTA CCAGGACGTC ACTGTCTACC TGGGAGACAC
CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC CAAATTTCCT
GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC

CGTGGAGAGC CGCATCACCC TGCACGAAAA CCGGACCCTT TCCATCAAGG
AGGCGTCCTT CTCAGACAGA GGCGTCTATA AGTGCGTGGC CAGCAATGCA
GCCGGGGCGG ACAGCCTGGC CATCCGCCTG CACGTGGCGG CACTGCCCCC
CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC CCGGGGCTCA
GCATTCACAT TCACTGCACT GCCAAGGCTG CGCCCCTGCC CAGCGTGCGC
TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG
GAACTTGTTT GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC
CCAAGGACAG CGGGCGCTAT GAGTGCGTGG CCGCCAACCT GGTAGGCTCC
GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCGCG
CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC
TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG
AATCAAGGTG TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA
AAGATGCCGG AGATTACCTG TGCGTAGCTC GAAATAAGGT TGGTGATGAC
TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA
CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC CTGAAAGTGG
ACTGTGTGGC CACCGGGCTT CCCAATCCCG AGATCTCCTG GAGCCTCCCA
GACGGGAGTC TGGTGAACTC CTTCATGCAG TCGGATGACA GCGGTGGACG
CACCAAGCGC TATGTCGTCT TCAACAATGG GACACTCTAC TTTAACGAAG
TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTTGCTGA AAATCAGGTC
GGGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC
CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG
TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT
ATACCAAGAT GGCACTCTCC TTATTCAGAA AGCCCAGCGT TCTGACAGCG
GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG
GTGTGGATTC ACGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA
CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAAACTGA
TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT
CCCGAGGGTG TGGTTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT
CCATGGCAAC GGTTCCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG
TCCAGCTGGT ATGCATGGCA CGCAACGAGG GAGGGGAGGC GAGGTTGATC
GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC
GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT
GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT
GGCACCGATC TGCAGAGTGG ACAGCAGCTG CAGCGCTTCT ACCACAAGGC
TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT
ACCGCTGCGT GGCCCGCAAT GCCGCTGGCC ACACGGAGAG GCTGGTCTCC
CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT
CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG
CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG

GGCCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC
GGTTCGTGAG GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG
AGACGGAGTA CGGCCCTTCG GTCACCAGCA TCCCCGTGAT TGTGATCGCC
TATCCTCCCC GGATCACCAG CGAGCCCACC CCGGTCATCT ACACCCGGCC
CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT CCCAAAGCTG
ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTCAG
GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCCAGGGAT CACTGACCAT
CCAGCATGCC ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA
ACATTCTCGG CAGTGACTCC AAAACAACTT ACATCCACGT CTTCTGAAAT
GTGGATTCCA GAATGATTGC TTAGGAACTG ACAACAAAGC GGGGTTTGTA
AGGGAAGCCA GGTTGGGGAA TAGGAGCTCT TAAATAATGT GTCACAGTGC
ATGGTGGCCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT
GAGACACTTT CTTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG
TCTGTGCTCT GACTGCAATT TTTCTTCTTT TGCAAATGCC ACTCGACTGC
CTTCATAAGC GTCCATAGGA TATCTGAGGA ACATTCATCA AAAATAAGCC
ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA TCACCTAGTT
AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCCAGAT TGACAAGTCA
TCTTTCAGTT ATTTCCTCTG TCACTTCAAA ACTCCAGCTT GCCCAATAAG
GATTTAGAAC CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT
TACATACATA CAGCTACCAT TTTATATGAA AAAAGAAAAA CATTTCTTCC
TGGAACTCAC TTTTTATATA ATGTTTTATA TATATATTTT TTCCTTTCAA
ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT ATTAAAATTA
ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA
ATATAATTTT AAAAAATTTC TCTCCAACCT CCTTCAAATT CAGTCACCAC
TGTTATATTA CCTTCTCCAG GAACCCTCCA GTGGGGAAGG CTGCGATATT
AGATTTCCTT GTATGCAAAG TTTTTGTTGA AAGCTGTGCT CAGAGGAGGT
GAGAGGAGAG GAAGGAGAAA ACTGCATCAT AACTTTACAG AATTGAATCT
AGAGTCTTCC CCGAAAAGCC CAGAAACTTC TCTGCAGTAT CTGGCTTGTC
CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC
CCATGAATAA TACACGACCT GTTATTTCCA TGACTGCTTT ACTGTATTTT
TAAGGTCAAT ATACTGTACA TTTGATAATA AAATAATATT CTCCCAAAAA AAAAA

Figure 53 **Human OCP: nucleotide sequence of ORF**

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGC
CTGGTCGCCACCCCTGGGGGCAAGGCCTGTCCTCGCCGCTGTGCCTGTTATATGCCT
ACGGAGGTACACTGCACATTTCGGTACCTGACTTCCATCCCAGACAGCATCCCGCCC
AATGTGGAACGCATCAATTTAGGATACAACAGCTTGGTTAGATTGATGGAAACAGAT
TTTTCTGGCCTGACCAAACTGGAGTTACTCATGCTTCACAGCAATGGCATTCACACA
ATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTAT
AATAAAGTCCGAAAACTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGA
TTGCACATGGACCACAACAATATTGAGTTTATAAACCCAGAGGTTTTTTATGGGCTC
AACTTTCTCCGCCTGGTGCACTTGGAAGGAAATCAGCTCACTAAGCTCCACCCAGAT
ACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAAATCTCTTTCATTAAGTTCCTA
TACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTATATGCCT
GACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAG
TGGTTGTCTGACTGGATACAGGAGAAGCCAGATGTAATAAAATGCAAAAAAGATAGA
AGTCCCTCTAGTGCTCAGCAGTGTCCACTTTGCATGAACCCTAGGACTTCTAAAGGC
AAGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAGTGTGCCAAGCCAACCATTGAC
TCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTTCATC
TCTCCCCAAGGTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGATCAG
TCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAGGACATCACCC
ATTGCATTCACTGAAGAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTT
TTGGTGTGCAACATAGATTACGGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTG
TACAGTGATTCTCCTCTGATACTAGAAAGGAGCCACTTGCTTAGTGAAACACCGCAG
CTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATTTTTACCAACATA
GAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCCTTGCAG
CTGAACAGAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCT
CAAATCACTTTACCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATT
TCAAGGGATAACAATACTAAGCTGGAACATACTGTCTTGGTAGGTGGAACCGTTGGC
CTGAACTGCCCAGGCCAAGGAGACCCCACCCCACACGTGGATTGGCTTCTAGCTGAT
GGAAGTAAAGTGAGAGCCCCTTATGTCAGTGAGGATGGACGGATCCTAATAGACAAA
AGTGGAAAATTGGAACTCCAGATGGCTGATAGTTTTGACACAGGCGTATATCACTGT
ATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAA
CCTTTGGTCGAAGCCTATCAGGAAAATGGGATTCATCACACAGTTTTCATTGGTGAA
ACACTTGATCTTCCATGCCATTCTACTGGTATCCCAGATGCCTCTATTAGCTGGGTT
ATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAAGTTCTAAACAAT
GGCACATTAAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGCTGTGTG
GCAGCCAACCCATCAGGGGTTGATTTTTTGATTTTCCAAGTTTCAGTCAAGATGAAA
GGACAAAGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCC
AATCCTATTGCTCATCTTAAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTG
ATGGAGGCTGAGGTTGGAAAACACACCTCAAGCACAAGTAAGAGGCACAACTATCGG
GAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGAATAGG
AGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCACTGTTG
GAGAAAGCTAAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCA
CCCCCAGTGGTCACCCAACTCCCAAACATACCTGGTGAAGAAGACGATTCCTCAGGC
ATGCTCGCTCTACATGAGGAATTTATGGTCCCGGCCACTAAAGCTTTGAACCTTCCA
GCAAGGACAGTGACTGCTGACTCCAGAACAATATCTGATAGTCCTATGACAAACATA
AATTATGGCACAGAATTCTCTCCTGTTGTGAATTCACAAATACTACCACCTGAAGAA
CCCACAGATTTCAAACTGTCTACTGCTATTAAAACTACAGCCATGTCAAAGAATATA

AACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCATCCACTGTC
TTTCCACTGCTACTTGGAGCAACTGAATTTCAGGACTCTGACCAGATGGGAAGAGGA
AGAGAGCATTTCCAAAGTAGACCCCCAATAACAGTAAGGACTATGATCAAAGATGTC
AATGTCAAAATGCTTAGTAGCACCACCAACAAACTATTATTAGAGTCAGTAAATACC
ACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCACTTC
TATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACA
GCTGCTCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTTC
AGACGCTTTGGGAGGCAGAGGAAAATTGGCGGAAGGGGGCGGATTATCAGCCCATAT
AGAACTCCAGTTCTGCGACGGCATAGATACAGCATTTTCAGGTCAACAACCAGAGGT
TCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAATGTGACATGTCTG
TCCTGTCTTCCCAGGGAGAGGCTCACCACTGCCACAGCAGCATTGTCTTTTCCAAGT
GCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCT
ACAACTCTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGAGAAA
ACAACACCCACAATAAAATATTTCAGGACTGAAATTTCCCAAGTGACTCCAACTGGT
GCAGTCATGACATATGCTCCAACATCCATACCCATGGAAAAAACTCACAAAGTAAAC
GCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGATTCAGTGATTACA
TCGTCACTTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACAAGG
TTTTCAAGAAGGAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAA
GGCAGATTAAGGAATCAACATAAAGTTAGTTTACAAAAAAGCACAGCTGTGATGCTT
CCTAAAACATCTCCTGCTTTACCACAGAGACAAAGTTCCCCTTTCCATTTCACCACA
CTTTCAACAAGTGTGATGCAAATTCCATCTAATACCTTGACTACCGCTCACCACACT
ACGACCAAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCCTTCCCA
CCCCTTAACCCTATGCTTCCTAGTATTATAAGCAAAGACTCAAGTACAAAAAGCATC
ATATCAACGCAAACAGCAATACCAGCAACAACTCCTACCTTCCCTGCATCTGTCATC
ACTTATGAAACCCAAACAGAGAGATCTAGAGCACAAACAATACAAAGAGAACAGGAG
CCTCAAAAGAAGAACAGGACTGACCCAAACATCTCTCCAGACCAGAGTTCTGGCTTC
ACTACACCCACTGCTATGACACCTCCTGCTCTGGCATTCACTCATTCCCCACCAGAA
AACACAACTGGGATTTCAAGCACAATCAGTTTTCATTCAAGAACTCTTAATCTGACA
GATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCT
TCTGAAACAACTTTGTCCAGCAAATCACACCAGAGTACCACAACTAGGAAAGCATCA
TTAGACACTCCCATACCACCATTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCC
ATCTCCCCTCCCTTTACTCAGAGAGCAGTTACTGACACACGTGGCGACTCCCATTTC
CGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGGCACAATCTCCAA
ATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTACATTCT
ACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAACTCCATCT
CCCTGGGCAGAATACCAATTTTGGCACAAACCATACTCAGACATTGCTGAAAAAGGC
AAAAAGCCAGAAGTAAGCATGTTGGCTACTACAGGCCTGTCCGAGGCCACCACTCTT
GTTTCAGATTGGGATGGACAGAAGAACACAAAGAAGAGTGACTTTGATAAGAAACCA
GTTCAAGAAGCAACAACTTCCAAACTCCTTCCCTTTGACTCTTTGTCTAGGTATATA
TTTGAAAAGCCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTTACTATTCCAGCTAAC
TCAGATGCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCCTGCCCACCATTCATTGG
ACCAGAGTTTCAGGACTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTT
CTCCCCAATGGTACCCTGTCCATCCAGAGGGTGGAAATTCAGGACCGCGGACAGTAC
TTGTGTTCCGCATCCAATCTGTTTGGCACAGACCACCTTCATGTCACCTTGTCTGTG
GTTTCCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTTCATTCC
GGAAGCACTGTGGAACTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACC
TGGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCT
GTGGTGACGGTTGACGGAACATTGGTCCTCCACAATCTCAGTATTTATGACCGTGGC
TTTTACAAATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGCTGGTTAAAATA
CAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGC
ACTTGGGGTGAAAGTTTAAAACTGCCCTGTACTGCAAAAGGAACTCCTCAGCCCAGC

GTTTACTGGGTCCTCTCTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCC
AAGTTGTTCTTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGAC
AGGGGCACTTATGAATGCATTGCTACCAGTTCCACTGGTTCGGAGCGAAGAGTAGTA
ATGCTTACAATGGAAGAGCGAGTGACCAGCCCCAGGATAGAAGCTGCATCCCAGAAA
AGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAACTGCTCAGCCACTGGGGAG
CCCAAACCCCAAATAATGTGGAGGTTACCATCCAAGGCTGTGGTCGACCAGTGGAGC
TGGATCCACGTCTACCCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGAC
AGTGGTGTCTACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATG
CATGTTAGCCTAAGACTGAAACCTGCCAAAATTGACCACAAGCAGTATTTTAGAAAG
CAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGCAAAGCTTCCGGCTCCCCAGTG
CCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGATCAACAATGCAATGCAAGCC
GATGACAGTGGCCACAGGACTAGGAGATATACCCTTTTCAACAATGGAACTTTATAC
TTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTTGCTATGCCCAGAACACC
CTAGGGAAAGATGAAATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATA
AGGCAGAGTAACAAAACCAACAAGAGAATCAAAGCTGGAGACACAGCTGTCCTTGAC
TGTGAGGTCACTGGGGATCCCAAACCAAAAATATTTTGGTTGCTGCCTTCCAATGAC
ATGATTTCCTTCTCCATTGATAGGTACACATTTCATGCCAATGGGTCTTTGACCATC
AACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCAGT
GGGGATGACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAAT
GGTCTGTATACAAACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAA
CACTTTGACTGCAGAGCTGAAGGGACACCATCTCCTGAAGTCATGTGGATCATGCCA
GACAATATTTTCCTCACAGCCCCATACTATGGAAGCAGAATCACAGTCCATAAAAAT
GGAACCTTGGAAATTAGGAATGTGAGGCTTTCAGATTCAGCCGACTTTATCTGTGTG
GCCCGAAATGAAGGTGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTACTGGAAATG
CTGAGAAGACCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCCAGCTGGGA
AAGTCCACAGCATTGAATTGCTCTGTTGATGGTAACCCACCACCTGAAATAATCTGG
ATTTTACCAAATGGCACACGATTTTCCAATGGACCACAAAGTTATCAGTATCTGATA
GCAAGCAATGGTTCTTTTATCATTTCTAAAACAACTCGGGAGGATGCAGGAAAATAT
CGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATT
GGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGA
GAATCTCTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGG
ACTATGCCAAGTGGTTATGTAGTAGACAGGCCTCAAATTAATGGGAAATACATATTG
CATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCTTATGACAGAGGAAACTAT
ATCTGTAAGGCTCAAAATAGTGTTGGTCATACACTGATTACTGTTCCAGTAATGATT
GTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGGACA
GGGGCAGCCTTTCAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAGAAATCACA
TGGGAGATGCCTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGA
AGTGAGCAGCTTCACTTACAAGGTACCCTAGTCATTCAGAATCCCCAAACCTCCGAT
TCTGGGATATACAAATGCACAGCAAAGAACCCACTTGGTAGTGATTATGCAGCAACG
TATATTCAAGTAATCTGA

Figure 54 **Human OCP: predicted amino acid sequence**

MKVKGRGITC LLVSFAVICL VATPGGKACP RRCACYMPTE VHCTFRYLTS
IPDSIPPNVE RINLGYNSLV RLMETDFSGL TKLELLMLHS NGIHTIPDKT
FSDLQALQVL KMSYNKVRKL QKDTFYGLRS LTRLHMDHNN IEFINPEVFY
GLNFLRLVHL EGNQLTKLHP DTFVSLSYLQ IFKISFIKFL YLSDNFLTSL
PQEMVSYMPD LDSLYLHGNP WTCDCHLKWL SDWIQEKPDV IKCKKDRSPS
SAQQCPLCMN PRTSKGKPLA MVSAAAFQCA KPTIDSSLKS KSLTILEDSS
SAFISPQGFM APFGSLTLNM TDQSGNEANM VCSIQKPSRT SPIAFTEEND
YIVLNTSFST FLVCNIDYGH IQPVWQILAL YSDSPLILER SHLLSETPQL
YYKYKQVAPK PEDIFTNIEA DLRADPSWLM QDQISLQLNR TATTFSTLQI
QYSSDAQITL PRAEMRPVKH KWTMISRDNN TKLEHTVLVG GTVGLNCPGQ
GDPTPHVDWL LADGSKVRAP YVSEDGRILI DKSGKLELQM ADSFDTGVYH
CISSNYDDAD ILTYRITVVE PLVEAYQENG IHHTVFIGET LDLPCHSTGI
PDASISWVIP GNNVLYQSSR DKKVLNNGTL RILQVTPKDQ GYYRCVAANP
SGVDFLIFQV SVKMKGQRPL EHDGETEGSG LDESNPIAHL KEPPGAQLRT
SALMEAEVGK HTSSTSKRHN YRELTLQRRG DSTHRRFREN RRHFPPSARR
IDPQHWAALL EKAKKNAMPD KRENTTVSPP PVVTQLPNIP GEEDDSSGML
ALHEEFMVPA TKALNLPART VTADSRTISD SPMTNINYGT EFSPVVNSQI
LPPEEEPTDFK LSTAIKTTAM SKNINPTMSS QIQGTTNQHS STVFPLLLGA
TEFQDSDQMG RGREHFQSRP PITVRTMIKD VNVKMLSSTT NKLLLESVNT
TNSHQTSVRE VSEPRHNHFY SHTTQILSTS TFPSDPHTAA HSQFPIPRNS
TVNIPLFRRF GRQRKIGGRG RIISPYRTPV LRRHRYSIFR STTRGSSEKS
TTAFSATVLN VTCLSCLPRE RLTTATAALS FPSAAPITFP KADIARVPSE
ESTTLVQNPL LLLENKPSVE KTTPTIKYFR TEISQVTPTG AVMTYAPTSI
PMEKTHKVNA SYPRVSSTNE AKRDSVITSS LSGAITKPPM TIIAITRFSR
RKIPWQQNFV NNHNPKGRLR NQHKVSLQKS TAVMLPKTSP ALPQRQSSPF
HFTTLSTSVM QIPSNTLTTA HHTTTKTHNP GSLPTKKELP FPPLNPMLPS
IISKDSSTKS IISTQTAIPA TTPTFPASVI TYETQTERSR AQTIQREQEP
QKKNRTDPNI SPDQSSGFTT PTAMTPPALA FTHSPPENTT GISSTISFHS
RTLNLTDVIE ELAQASTQTL KSTIASETTL SSKSHQSTTT RKASLDTPIP
PFLSSSATLM PVPISPPFTQ RAVTDTRGDS HFRLMTNTVV KLHESSRHNL
QMPSSQLEPL TSSTSNLLHS TPMPALTTVK SQNSKLTPSP WAEYQFWHKP
YSDIAEKGKK PEVSMLATTG LSEATTLVSD WDGQKNTKKS DFDKKPVQEA
TTSKLLPFDS LSRYIFEKPR IVGGKAASFT IPANSDAFLP CEAVGNPLPT
IHWTRVSGLD LSRGNQNSRV QVLPNGTLSI QRVEIQDRGQ YLCSASNLFG
TDHLHVTLSV VSYPPRILER RTKEITVHSG STVELKCRAE GRPSPTVTWI
LANQTVVSES SQGSRQAVVT VDGTLVLHNL SIYDRGFYKC VASNPGGQDS
LLVKIQVIAA PPVILEQRRQ VIVGTWGESL KLPCTAKGTP QPSVYWVLSD
GTEVKPLQFT NSKLFLFSNG TLYIRNLASS DRGTYECIAT SSTGSERRVV
MLTMEERVTS PRIEAASQKR TEVNFGDKLL LNCSATGEPK PQIMWRLPSK
AVVDQWSWIH VYPNGSLFIG SVTEKDSGVY LCVARNKMGD DLILMHVSLR
LKPAKIDHKQ YFRKQVLHGK DFQVDCKASG SPVPEISWSL PDGTMINNAM
QADDSGHRTR RYTLFNNGTL YFNKVGVAEE GDYTCYAQNT LGKDEMKVHL
TVITAAPRIR QSNKTNKRIK AGDTAVLDCE VTGDPKPKIF WLLPSNDMIS
FSIDRYTFHA NGSLTINKVK LLDSGEYVCV ARNPSGDDTK MYKLDVVSKP
PLINGLYTNR TVIKATAVRH SKKHFDCRAE GTPSPEVMWI MPDNIFLTAP
YYGSRITVHK NGTLEIRNVR LSDSADFICV ARNEGGESVL VVQLEVLEML
RRPTFRNPFN EKIVAQLGKS TALNCSVDGN PPPEIIWILP NGTRFSNGPQ
SYQYLIASNG SFIISKTTRE DAGKYRCAAR NKVGYIEKLV ILEIGQKPVI

LTYAPGTVKG ISGESLSLHC VSDGIPKPNI KWTMPSGYVV DRPQINGKYI
LHDNGTLVIK EATAYDRGNY ICKAQNSVGH TLITVPVMIV AYPPRITNRP
PRSIVTRTGA AFQLHCVALG VPKPEITWEM PDHSLLSTAS KERTHGSEQL
HLQGTLVIQN PQTSDSGIYK CTAKNPLGSD YAATYIQVI*